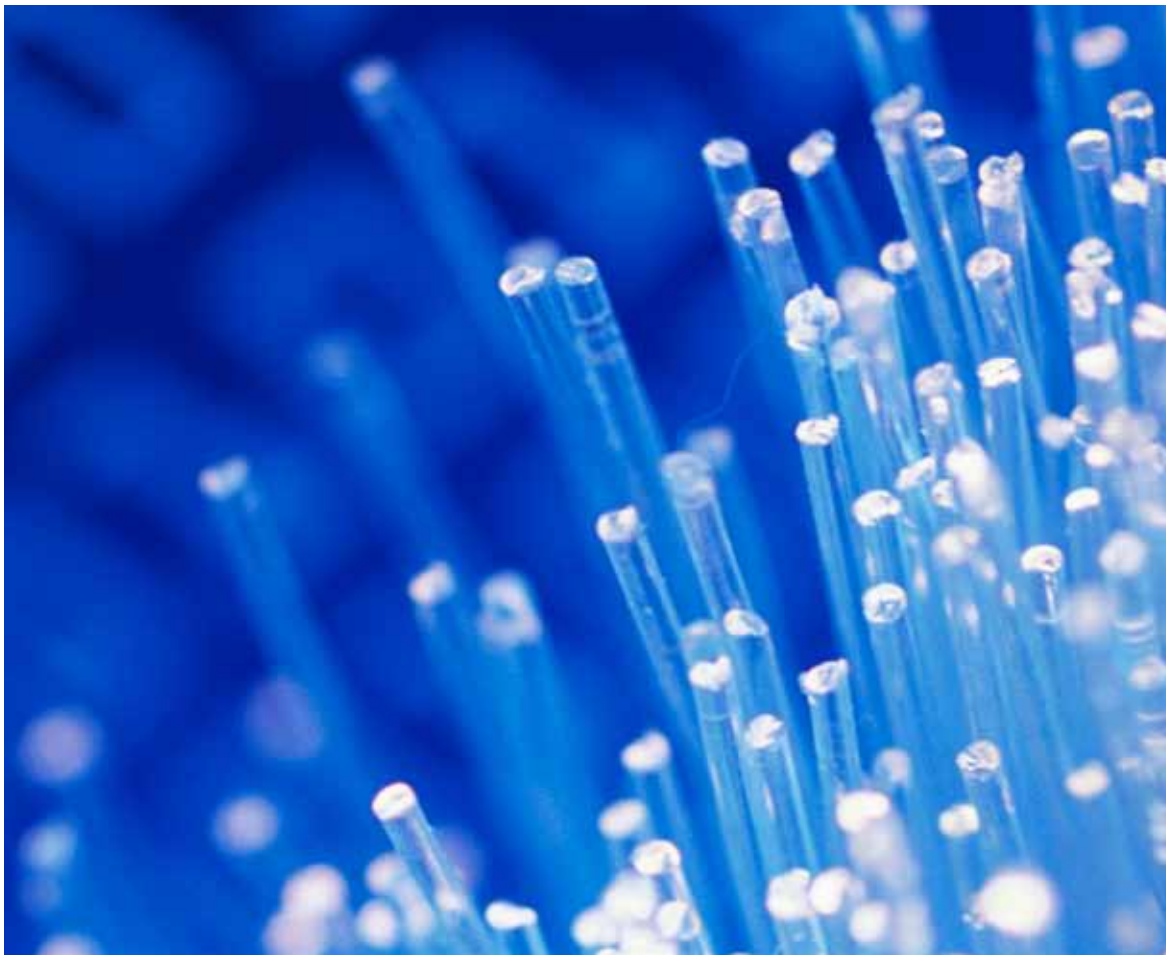
A Roadmap for Cybersecurity Research



Homeland Security

November 2009

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
NOV 2009		**00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Roadmap for Cybersecurity Research	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Homeland Security,Washington,DC**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**126**	

Contents

Executive Summary

Executive Summary

The United States is at a significant decision point. We must continue to defend our current systems and networks and at the same time attempt to "get out in front" of our adversaries and ensure that future generations of technology will position us to better protect our critical infrastructures and respond to attacks from our adversaries. The term "system" is used broadly to encompass systems of systems and networks.

This cybersecurity research roadmap is an attempt to begin to define a national R&D agenda that is required to enable us to get ahead of our adversaries and produce the technologies that will protect our information systems and networks into the future. The research, development, test, evaluation, and other life cycle considerations required are far reaching—from technologies that secure individuals and their information to technologies that will ensure that our critical infrastructures are much more resilient. The R&D investments recommended in this roadmap must tackle the vulnerabilities of today and envision those of the future.

The intent of this document is to provide detailed research and development agendas for the future relating to 11 hard problem areas in cybersecurity, for use by agencies of the U.S. Government and other potential R&D funding sources. The 11 hard problems are:

1. Scalable trustworthy systems (including system architectures and requisite development methodology)
2. Enterprise-level metrics (including measures of overall system trustworthiness)
3. System evaluation life cycle (including approaches for sufficient assurance)
4. Combatting insider threats
5. Combatting malware and botnets
6. Global-scale identity management
7. Survivability of time-critical systems
8. Situational understanding and attack attribution
9. Provenance (relating to information, systems, and hardware)
10. Privacy-aware security
11. Usable security

For each of these hard problems, the roadmap identifies critical needs, gaps in research, and research agenda appropriate for near, medium, and long term attention.

DHS S&T assembled a large team of subject matter experts who provided input into the development of this research roadmap. The content was developed over the course of 15 months that included three regional multi-day workshops, two virtual workshops for each topic, and numerous editing activities by the participants.

Introduction

Introduction

Information technology has become pervasive in every way—from our phones and other small devices to our enterprise networks to the infrastructure that runs our economy. Improvements to the security of this information technology are essential for our future. As the critical infrastructures of the United States have become more and more dependent on public and private networks, the potential for widespread national impact resulting from disruption or failure of these networks has also increased. Securing the nation's critical infrastructures requires protecting not only their physical systems but, just as important, the cyber portions of the systems on which they rely. The most significant cyber threats to the nation are fundamentally different from those posed by the "script kiddies" or virus writers who traditionally have plagued users of the Internet. Today, the Internet has a significant role in enabling the communications, monitoring, operations, and business systems underlying many of the nation's critical infrastructures. Cyberattacks are increasing in frequency and impact. Adversaries seeking to disrupt the nation's critical infrastructures are driven by different motives and view cyberspace as a possible means to have much greater impact, such as causing harm to people or widespread economic damage. Although to date no cyberattack has had a significant impact on our nation's critical infrastructures, previous attacks have demonstrated that extensive vulnerabilities exist in information systems and networks, with the potential for serious damage. The effects of a successful attack might include serious economic consequences through impacts on major economic and industrial sectors, threats to infrastructure elements such as electric power, and disruptions that impede the response and communication capabilities of first responders in crisis situations.

The United States is at a significant decision point. We must continue to defend our current systems and networks and at the same time attempt to "get out in front" of our adversaries and ensure that future generations of technology will position us to better protect our critical infrastructures and respond to attacks from our adversaries. It is the opinion of those involved in creating this research roadmap that government-funded research and development (R&D) must play an increasing role to enable us to accomplish this goal of national and economic security. The research topics in this roadmap, however, are relevant not only to the federal government but also to the private sector and others who are interested in securing the future.

This cybersecurity research roadmap is an attempt to begin to define a national R&D agenda that is required to enable us to get ahead of our adversaries and produce the technologies that will protect our information systems and networks into the future. The research, development, test, evaluation, and other life cycle considerations required are far reaching—from technologies that secure individuals and their information to technologies that will ensure that our critical infrastructures are much more resilient. These investments must tackle the vulnerabilities of today and envision those of the future.

"The time is now near at hand..."
— George Washington, July 2, 1776

Historical background

The INFOSEC Research Council (IRC) is an informal organization of government program managers who sponsor information security research within the U.S. Government. Many organizations have representatives as regular members of the IRC: Central Intelligence Agency, Department of Defense (including the Air Force, Army, Defense Advanced Research Projects Agency, National Reconnaissance Office, National Security Agency, Navy, and Office of the Secretary of Defense), Department of Energy, Department of Homeland Security, Federal Aviation Administration, Intelligence Advanced Research Projects Activity, National Aeronautics and Space Administration, National Institutes of Health, National Institute of Standards and Technology, National Science Foundation, and the Technical Support Working Group. In addition, the IRC is regularly attended by partner organizations from Canada and the United Kingdom.

The IRC developed the original Hard Problem List (HPL), which was composed in 1997 and published in draft form in 1999. The HPL defines desirable research topics by identifying a set of key problems from the U.S. Government perspective and in the context of IRC member missions. Solutions to these problems would remove major barriers to effective information security (INFOSEC). The Hard Problem List was intended to help guide the research program planning of the IRC member organizations. It was also hoped that nonmember organizations and industrial partners would consider these problems in the development of their research programs. The original list has proven useful in guiding INFOSEC research, and policy makers and planners may find the document useful in evaluating the contributions of ongoing and proposed INFOSEC research programs. However, the significant evolution of technology and threats between 1999 and 2005 required an update to the list. Therefore, an updated version of the HPL was published in November 2005. This updated document included the following technical hard problems from the information security perspective:

1. Global-Scale Identity Management
2. Insider Threat
3. Availability of Time-Critical Systems
4. Building Scalable Secure Systems
5. Situational Understanding and Attack Attribution
6. Information Provenance
7. Security with Privacy
8. Enterprise-Level Security Metrics

These eight problems were selected as the hardest and most critical challenges that must be addressed by the INFOSEC research community if trustworthy systems envisioned by the U.S. Government are to be built. INFOSEC problems may be characterized as "hard" for several reasons. Some problems are hard because of the fundamental technical challenges of building secure systems, others because of the complexity of information technology (IT) system applications. Contributing to these problems are conflicting regulatory and policy goals, poor understanding of operational needs and user interfaces, rapid changes in technology, large heterogeneous environments (including mixes of legacy systems), and the presence of significant, asymmetric threats.

The area of cybersecurity and the associated research and development activities have been written about frequently over the past decade. In addition to both the original IRC HPL in 1999 and the revision in 2005, the following reports have discussed the need for investment in this critical area:

- Toward a Safer and More Secure Cyberspace
- Federal Plan for Cyber Security and Information Assurance Research and Development
- Cyber Security: A Crisis of Prioritization
- Hardening the Internet
- Information Security Governance: A Call to Action
- The National Strategy to Secure Cyberspace
- Cyber Security Research and Development Agenda

These reports can be found at http://www.cyber.st.dhs.gov/documents.html

Current context

On January 8, 2008, the President issued National Security Presidential Directive 54/Homeland Security Presidential Directive 23, which formalized the Comprehensive National Cybersecurity Initiative (CNCI) and a series of continuous efforts designed to establish a frontline defense (reducing current vulnerabilities and preventing intrusions), defending against the full spectrum of threats by using intelligence

and strengthening supply chain security, and shaping the future environment by enhancing our research, development, and education, as well as investing in "leap-ahead" technologies.

The vision of the CNCI research community over the next 10 years is to "transform the cyber-infrastructure so that critical national interests are protected from catastrophic damage and our society can confidently adopt new technological advances."

Two components of the CNCI deal with cybersecurity research and development—one focused on the coordination of federal R&D and the other on the development of leap-ahead technologies.

No single federal agency "owns" the issue of cybersecurity. In fact, the federal government does not uniquely own cybersecurity. It is a national and global challenge with far-reaching consequences that requires a cooperative, comprehensive effort across the public and private sectors. However, as it has done historically, U.S. Government R&D in key technologies working in close cooperation with private-sector partners can jump-start the necessary fundamental technical transformation.

The leap-ahead strategy aligns with the consensus of the nation's networking and cybersecurity research communities that the only long-term solution to the vulnerabilities of today's networking and information technologies is to ensure that future generations of these technologies are designed with security built in from the ground up. The leap-ahead strategy will help extend U.S. leadership at a time of growing influence in networking and IT systems, components, and standards among U.S. competitors. Federal agencies with mission-critical needs for increased cybersecurity, which includes information assurance as well as network and system security, can play a direct role in determining research priorities and assessing emerging technology prototypes. Moreover, through technology transfer efforts, the federal government can encourage rapid adoption of the results of leap-ahead research. Technology breakthroughs that can curb or break the resource-draining cycle of security patching will have a high likelihood of marketplace implementation.

As stated previously, this Cybersecurity Research Roadmap is an attempt to begin to address a national R&D agenda that is required to enable us to get ahead of our adversaries and produce the technologies that will protect our information systems and networks into the future. The topics contained in this roadmap and the research and development that would be accomplished if the roadmap were implemented are, in fact, leap-ahead in nature and address many of the topics that have been identified in the CNCI activities

Document format

The intent of this document is to provide detailed research and development agendas for the future relating to 11 hard problem areas in cybersecurity, for use by agencies of the U.S. Government and anyone else that is funding or doing R&D. It is expected that each agency will find certain parts of the document resonant with its own needs and will proceed accordingly with some interagency coordination to ensure coverage of all the topics.

Each of the following topic areas is treated in detail in a subsequent section of its own, from Section 1 to Section 11.

1. Scalable trustworthy systems (including system architectures and requisite development methodology)
2. Enterprise-level metrics (including measures of overall system trustworthiness)
3. System evaluation life cycle (including approaches for sufficient assurance)
4. Combatting insider threats
5. Combatting malware and botnets
6. Global-scale identity management
7. Survivability of time-critical systems
8. Situational understanding and attack attribution
9. Provenance (relating to information, systems, and hardware)
10. Privacy-aware security
11. Usable security

Eight of these topics (1, 2, 4, 6, 7, 8, 9, 10) are adopted from the November 2005 IRC Hard Problem List [IRC05] and are still of vital relevance. The other three topics (3, 5, 11) represent additional areas considered to be of particular importance for the future.

The order in which the 11 topics are presented reflects some structural similarities among subgroups of the topics and exhibits clearly some of their major interdependencies. The order proceeds roughly from overarching system concepts to more detailed issues—except

for the last topic—and has the following structure:

a. Topics 1–3 frame the overarching problems.
b. Topics 4–5 relate to specific major threats and needs.
c. Topics 6–10 relate to some of the "ilities" and to system concepts required for implementing the previous topics.

Topic 11, usable security, is different from the others in its cross-cutting nature. If taken seriously enough, it can influence the success of almost all the other topics. However, some sort of transcendent usability requirements need to be embedded pervasively in all the other topics.

Each of the 11 sections follows a similar format. To get a full picture of the problem, where we are, and where we need to go, we ask the following questions:

Background

- What is the problem being addressed?
- What are the potential threats?
- Who are the potential beneficiaries? What are their respective needs?
- What is the current state of the practice?
- What is the status of current research?

Future Directions

- On what categories can we subdivide the topics?
- What are the major research gaps?
- What are some exemplary problems for R&D on this topic?
- What are the challenges that must be addressed?
- What approaches might be desirable?
- What R&D is evolutionary and what is more basic, higher risk, game changing?
- Resources
- Measures of success
- What needs to be in place for test and evaluation?
- To what extent can we test real systems?

Following the 11 sections are three appendices:

Appendix A: Interdependencies among Topics

Appendix B: Technology Transfer

Appendix C: List of Participants in the Roadmap Development

References

[IRC2005] INFOSEC Research Council Hard Problem List, November 2005 http://www.cyber.st.dhs.gov/docs/IRC_Hard_Problem_List.pdf.

[USAF-SAB07] United States Air Force Scientific Advisory Board, Report on Implications of Cyber Warfare. Volume 1: Executive Summary and Annotated Brief; Volume 2: Final Report, August 2007. For Official Use Only.

Additional background documents (including the two most recent National Research Council study reports on cybersecurity) can be found online. (http://www.cyber.st.dhs.gov/documents.html).

Acknowledgements

Acknowledgements

The content of this research roadmap was developed over the course of 15 months that included three workshops, two phone sessions for each topic, and numerous editing activities by the participants. Appendix C lists all the participants. The Cyber Security program of the Department of Homeland Security (DHS) Science and Technology (S&T) Directorate would like to express its appreciation for the considerable amount of time they dedicated to this effort.

DHS S&T would also like to acknowledge the support provided by the staff of SRI International in Menlo Park, CA, and Washington, DC. SRI is under contract with DHS S&T to provide technical, management, and subject matter expert support for the DHS S&T Cyber Security program. Those involved in this effort include Gary Bridges, Steve Dawson, Drew Dean, Jeremy Epstein, Pat Lincoln, Ulf Lindqvist, Jenny McNeill, Peter Neumann, Robin Roy, Zach Tudor, and Alfonso Valdes.

Of particular note is the work of Jenny McNeill and Peter Neumann. Jenny has been responsible for the organization of each of the workshops and phone sessions and has worked with SRI staff members Klaus Krause, Roxanne Jones, and Ascencion Villanueva to produce the final document. Peter Neumann has been relentless in his efforts to ensure that this research roadmap represents the real needs of the community and has worked with roadmap participants and government sponsors to produce a high-quality product.

Current Hard Problems in INFOSEC Research

1. Scalable Trustworthy Systems

BACKGROUND

What is the problem being addressed?

Trustworthiness is a multidimensional measure of the extent to which a system is likely to satisfy each of multiple aspects of each stated requirement for some desired combination of system integrity, system availability and survivability, data confidentiality, guaranteed real-time performance, accountability, attribution, usability, and other critical needs. Precise definitions of what trustworthiness means for these requirements and well-defined measures against which trustworthiness can be evaluated are fundamental precursors to developing and operating trustworthy systems. These precursors cut across everything related to scalable trustworthy systems. If what must be depended on does not perform according to its expectations, then whatever must depend on it may itself not be trustworthy. A trusted system is one that must be assumed to satisfy its requirements—whether or not it is actually trustworthy; indeed, it is a system whose failure in any way may compromise those requirements. Unfortunately, today's systems are typically not well suited for applications with critical trustworthiness requirements.

Scalability is the ability to satisfy given requirements as systems, networks, and systems of systems expand in functionality, capacity, complexity, and scope of trustworthiness requirements security, reliability, survivability, and improved real-time performance. Scalability must typically be addressed from the outset; experience shows that scalability usually cannot be retrofitted into systems for which it was not an original design goal. Scalable trustworthiness will be essential for many national- and world-scale systems, including those supporting critical infrastructures. Current methodologies for creating high-assurance systems do not scale to the size of today's—let alone tomorrow's—critical systems.

Composability is the ability to create systems and applications with predictably satisfactory behavior from components, subsystems, and other systems. To enhance scalability in complex distributed applications that must be trustworthy, high-assurance systems should be developed from a set of composable components and subsystems, each of which is itself suitably trustworthy, within a system architecture that inherently supports facile composability. Composition includes the ability to run software compatibly on different hardware, aided considerably by abstraction, operating systems, and suitable programming languages. However, we do not yet have a suitable set of trustworthy building blocks, composition methodologies, and analytic tools that would ensure that trustworthy systems could be developed as systems of other systems. In addition, requirements and evaluations should also compose accordingly. In the future, it will be vital that new systems can be incrementally added to a system of systems with some predictable confidence that the trustworthiness of the resulting systems of systems will not be weakened—or indeed that it may be strengthened.

Growing interconnectedness among existing systems results, in effect, in new composite systems at increasingly large scales. Existing hardware, operating system, networking, and application architectures do not adequately account for combined requirements for security, performance, and usability—confounding attempts to build trustworthy systems on them. As a result, today the security of a system of systems may be drastically less than that of most of its components.

In certain cases, it may be possible to build systems that are more trustworthy than some (or even most) of their components—for example, through constructive system design and meticulous attention to good software engineering practices. Techniques for building more trustworthy systems out of less trustworthy components have long been known and used in practice (e.g., summarized in [Neu2004], in the context of composability). For example, error-correcting codes can overcome unreliable communications and storage media, and encryption can be used to increase confidentiality and integrity despite insecure communication channels. These techniques are incomplete by themselves and generally ignore many security threats. They typically depend on the existence of some combination of trustworthy developers, trustworthy systems, trustworthy users, and trustworthy administrators, and their trustworthy embedding in those systems.

The primary focus of this topic area is scalability that preserves and enhances trustworthiness in real systems. The perceived order of importance for research and development in this topic area is as follows: (1) trustworthiness, (2) composability, and (3) scalability. Thus, the challenge addressed here is threefold: (a) to provide a sound basis for composability that can scale to the development of large and complex trustworthy systems; (b) to stimulate the development of the components, analysis tools, and testbeds required for that effort; and (c) to ensure that trustworthiness evaluations themselves can be composed.

What are the potential threats?

Threats to a system in operation include everything that can prevent critical applications from satisfying their intended requirements, including insider and outsider misuse, malware and other system subversions, software flaws, hardware malfunctions, human failures, physical damage, and environmental disruptions. Indeed, systems sometimes fail without any external provocation, as a result of design flaws, implementation bugs, misconfiguration, and system aging. Additional threats arise in the system acquisition and code distribution processes. Serious security problems have also resulted from discarded or stolen systems. For large-scale systems consisting of many independent installations (such as the Domain Name System, DNS), security updates must reach and be installed in all relevant components throughout the entire life cycle of the systems. This scope of updating has proven to be difficult to achieve.

Critical systems and their operating environments must be trustworthy despite a very wide range of adversities and adversaries. Historically, many system uses assumed the existence of a trustworthy computing base that would provide a suitable foundation for such computing. However, this assumption has not been justified. In the future, we must be able to develop scalable trustworthy systems effectively.

Who are the potential beneficiaries? What are their respective needs?

Large organizations in all sectors—for example, government, military, commercial, financial, and energy—suffer the consequences of using large-scale computing systems whose trustworthiness either is not assured or is potentially compromised because of costs that outweigh the perceived benefits. All stakeholders have requirements for confidentiality, integrity, and availability in their computing infrastructures, although the relative importance of these requirements varies by application. Achieving scalability and evolvability of systems without compromising trustworthiness is a major need. Typical customers include the following:

- Large-system developers (e.g., of operating systems, database management systems, national infrastructures such as the power grid)
- Application developers
- Microelectronics developers
- System integrators
- Large- and small-scale users
- Purveyors of potential exemplar applications for scalable trustworthiness

Several types of systems suggest the importance of being able to develop

scalable trustworthy systems. Examples include the following:

- Air traffic control systems
- Power grids
- Worldwide funds transfer systems
- Cellphone networks

Such systems need to be robust and capable of satisfying the perceived trustworthiness requirements. Outages in these systems can be extremely costly and dangerous. However, the extent to which the underlying concepts used to build these existing systems can continue to scale and also be responsive to more exacting trustworthiness requirements is unknown—especially in the face of increasing cyberthreats. The R&D must provide convincing arguments that they will scale appropriately. Exemplars of potential component systems might include the following:

- Trustworthy handheld multipurpose devices and other end-user devices
- Trustworthy special-purpose servers
- Embedded control systems that can be composed and used effectively
- Trustworthy networks
- Navigation systems, such as the Global Positioning Systems (GPS)

One or more such systems should be chosen for deeper study to develop better understanding of the approaches to scalable security developed in this program. In turn, the results of ongoing work on scalable trustworthiness should be applied to those and other exemplars.

What is the current state of the practice?

Hardware developers have recently made significant investments in specification, formal methods, configuration control, modeling, and prediction, partly in response to recognized problems, such as the Intel floating point flaw, and partly as a result of increased demonstrations of the effectiveness of those techniques.

The foundation for trustworthy scalable systems is established by the underlying hardware architecture. Adequate hardware protections are essential, and nearly all extant hardware architectures lack needed capabilities. Examples include fine-grain memory protection, inaccessible program control state, unmodifiable executable codes, fully granular access protections, and virtually mapped bus access by I/O and other adapter boards.

Although it might be appealing to try to apply those approaches to software, the issues of scalability suggest that additional approaches may be necessary. Numerous software-related failures have occurred (e.g., see [Neu1995]). In addition, techniques are needed to address how software/hardware interactions affect the overall trust level. Unfortunately, there is no existing mandate for significant investment during software system development to ensure scalable trustworthiness. Consequently, such efforts are generally not adequately addressed.

Diagnostic tools to detect software flaws on today's hardware architectures may be useful in the short run but are insufficient in the long run. Research is needed to establish the repertoire of architected hardware protections that are essential for system trustworthiness. It is unlikely that software alone can ever compensate fully for the lack of such hardware protections.

A possible implication is that the commercial off-the-shelf (COTS) systems in pervasive use today will never become sufficiently trustworthy. If that is indeed true, testing that implication should be identified as an activity and milestone in the recommended research agenda.

Convincing hardware manufacturers and software developers to provide and support needed hardware capabilities, of course, is a fundamental obstacle. The manufacturers' main motivations are least change and time to market. Until compelling research findings, legal consequences (e.g., financial liability for customer damages), and economic forces (e.g., purchase policies mandating the needed capabilities) are brought to bear, it seems unlikely that goals for the securing COTS and open source products can be realized.

What is the status of current research?

Over the past decade, significant computer security investments have been made in attempts to create greater assurance for existing applications and computer-based enterprises that are based predominantly on COTS components. Despite some progress, there are severe limits to this approach, and success has been meager at best, particularly with respect to trustworthiness, composability, and scalability.

The assurance attainable by incremental improvements on COTS products is fundamentally inadequate for critical applications.

Various research projects over the past half-century have been aimed at the challenge of designing and evaluating scalable trustworthy systems and networks, with some important research contributions with respect to both hardware and software. Some of these date back to the 1960s and 1970s, such as Multics, PSOS (the Provably Secure Operating System) and its formally based Hierarchical Development Methodology (HDM), the Blacker system as an early example of a virtual private network, the CLInc (Computational Logic, Inc.) stack, Gypsy, InaJo, Euclid, ML and other functional programming languages, and the verifying compiler, to name just a few. However, very few systems available today have taken serious advantage of such potentially far-reaching research efforts, or even the rather minimal guidance of Security Level 4 in FIPS 140-1. Also, the valued but inadequately observed 1975 security principles of Saltzer and Schroeder have recently been updated by Saltzer and Kaashoek [Sal+2009].

Some more recent efforts can also be cited here. For example, architectures exist or are contemplated for robust hardware that would inherently increase system trustworthiness by avoiding common vulnerabilities, including modernized capability-based architectures. In addition, the Trusted Computing Exemplar Project at the Naval Postgraduate School (http://cisr.nps.edu/projects/tcx.html) is intended to provide a working example of how trustworthy computing systems can be designed and built. It will make all elements of the constructive security process openly available. Recent advances in cryptography can also help, although some composability issues remain to be resolved as to how to embed those advances securely into marginally secure computer systems. Also, public key infrastructures (PKIs) are becoming more widely used and embedded in applications. However, many gaps remain in reusable requirements for trustworthiness, system architectures, software engineering practices, sound programming languages that avoid many of the characteristic flaws, and analysis tools that scale up to entire systems. Thoroughly worked examples of trustworthy systems are needed that can clearly demonstrate that well-conceived composability can enhance both trustworthiness and scalability. For example, each of the exemplars noted above would benefit greatly from the incorporation of scalable trustworthy systems.

At present, even for small systems, there exist very few examples of requirements, trustworthiness metrics, and operational systems that encompass a broad spectrum of trustworthiness with any generality. Furthermore, such requirements, metrics, and systems need to be composable and scalable into trustworthy systems of systems. However, a few examples exist for dedicated special-purpose systems, such as data diodes enforcing one-way communication paths and the Naval Research Laboratory Pump enabling trustworthy reading of information at lower levels of multilevel security.

In recent years, research has advanced significantly in formal methods applicable to software trustworthiness. That research is generally more applicable to new systems rather than to being retrofitted into existing systems. However, it needs to focus on attributes and subsystems for which it can be most effective, and must deal with complexity, scalability, hardware and software, and practical issues such as device drivers and excessive root privileges.

FUTURE DIRECTIONS

On what categories can we subdivide this topic?

For present purposes, different approaches to development of trustworthy scalable systems are associated with the following three roadmap categories. These categories are distinguished from one another roughly based on the extent to which they are able to reuse existing components.

1. Improving trustworthiness in existing systems. This incremental approach could entail augmenting relatively untrustworthy systems with some trustworthy components and enforcing operational constraints in attempts to achieve either trustworthy functions or systems with more clearly understood trust properties. Can we make existing systems significantly more trustworthy without wholesale replacement?

2. Clean-slate approaches. This entails building trustworthy primitives, composing them into trustworthy functions, and then verifying the overall trust level of the composite system. How much

better would this be? Would this enable solutions of problems that cannot be adequately addressed today, and for what requirements? Under what circumstances and for what requirements might this be possible? What new technologies, system architectures, and tools might be needed?

3. Operating successfully for given requirements despite the presence of partially untrusted environments. For example, existing computing systems might be viewed as "enemy territory" because they have been subject to unknown influences within the commercial supply chain and the overall life cycle (design, implementation, operations, maintenance, and decommissioning).

It is inherently impossible to control every aspect of the entire life cycle and the surrounding operational environments. For example, end-to-end cryptography enables communications over untrustworthy media—but does not address denial-of-service attacks en route or insider subversion at the endpoints.

The three categories are not intended to be mutually exclusive. For example, hybrid approaches can combine legacy systems from category 1 with incremental changes and significant advances from category 2. Indeed, hybrids among these three categories are not merely possible but quite likely. For example, approaches that begin with a clean-slate architecture could also incorporate some improvements of existing systems, and even allow some operations to take place in untrusted environments—if suitably encapsulated, confined, or otherwise controlled. A clean-slate approach tolerating an ongoing level of continuous compromise in its system components might also be viewed as a hybrid of categories 2 and 3. Further R&D is clearly required to determine the tradeoffs in cost-effectiveness, practicality, performance, usability, and relative trustworthiness attainable for any particular set of requirements. DARPA's IAMANET is a step in that direction.

An urgent need exists for R&D on incremental, clean-slate, and hybrids approaches. Trustworthiness issues may affect the development process and the resulting system performance. Adding functionality and concomitant complexity to achieve trustworthiness may be counterproductive, if not done constructively; it typically merely introduces new vulnerabilities. Trustworthiness must be designed in from the outset with complete specified requirements. Functionality and trustworthiness are inherently in conflict in the design process, and this conflict must be resolved before any implementation.

What are the major research gaps?

Research relating to composability has addressed some of the fundamental problems and underlying theory. For example, see [Neu2004] for a recent consideration of past work, current practice, and R&D directions that might be useful in the future. It contains numerous references to papers and reports on composability. It also considers a variety of techniques for compositions of subsystems that can increase trustworthiness, as well as system and network architectures and system development practices that can yield greater trustworthiness. See also [Can2001], which represents the beginning of work on the notion of universal composability applied to cryptography.

However, there are gaps in our understanding of composability as it relates to security, and to trustworthiness more generally, primarily because we lack precise specifications of most of the important requirements and desired properties. For example, we are often good at developing specific solutions to specific security problems, but we do not understand how to apply and combine these specific solutions to produce trustworthy systems. We lack methods for analyzing how even small changes to systems affect their trustworthiness. More broadly, we lack a good understanding of how to develop and maintain trustworthy systems comprehensively throughout the entire life cycle. We lack methods and tools for decomposing high-level trustworthiness goals into specific design requirements, capturing and specifying security requirements, analyzing security requirements, mapping higher-layer requirements into lower-layer ones, and verifying system trustworthiness properties. We do not understand how to combine systems in ways that ensure that the combination is more, rather than less, secure and resilient than its weakest components. We lack a detailed case history of past successes and failures in the development of trustworthy systems that could help us to elucidate principles and properties of trustworthy systems, both in an overarching sense and in specific application areas. We lack development tools and languages that could enable separation of functionality and trustworthiness

concerns for developers. For small systems, ad hoc solutions seldom suffice if they do not reflect such fundamental understanding of the problems. For the large-scale, highly complex systems of the future, we cannot expect to achieve adequate trustworthiness without deeper understanding, better tools, and more reliable evaluation methods—as well as composable building blocks and well-documented, worked examples of less complex systems.

The research directions can be partitioned into near-term, medium-term, and long-term opportunities. In general, the near-term approaches fall into the incremental category, and the long-term approaches fall into clean-slate and hybrid categories. However, the long-term approaches often have staged efforts that begin with near-term efforts. Also, the hybrid efforts tend to require longer-term schedules because some of them rely on near- and medium-term efforts.

Near term

- Development of prototype trustworthy systems in selected application and infrastructure domains
- Exploitation of cloud architectures and web-based applications
- Development of simulation environments for testing approaches to development of scalable trustworthy systems
- Intensive further research in composability
- Development of building blocks for composing trustworthy systems
- Well-defined composable specifications for requirements and components
- Realistic provable security properties for small-scale systems
- Urgent need for detailed worked examples
- Better understanding of the security properties of existing major components.

Medium term

- New hardware with well-understood trustworthiness properties
- Better operating systems and networking
- Better application architectures for trustworthy systems
- Isolation of legacy systems in trustworthy virtualization environments
- Continued research in composability, techniques for verifying the security properties of composed systems in terms of their specifications
- Urgent need for detailed realistic and practical worked examples.

Long term

- Tools for verifying trustworthiness of composite systems
- Techniques and tools for developing and maintaining trustworthy systems throughout the life cycle
- More extensive detailed worked examples.

Several threads could run through this timeline—for example, R&D relating to trustworthy isolation, separation, and virtualization in hardware and software; composability of designs and implementations; analyses that could greatly simplify evaluation of trustworthiness before putting applications into operation; robust architectures that provide self-testing, self-diagnosing, self-reconfiguring, compromise resilient, and automated remediation; and architectures that break the current asymmetric advantage for attackers (offense is cheaper than defense, at present). The emphasis needs to be on realistic, practical approaches to developing systems that are scalable, composable, and trustworthy.

The gaps in practice and R&D, approaches, and potential benefits are summarized in Table 1.1. The research directions for scalable trustworthy systems are intended to address these gaps. Table 1.2 also provides a summary of this section.

This topic area interacts strongly with enterprise-level metrics (Section 2) and evaluation methodology (Section 3) to provide assurance of trustworthiness. In the absence of such metrics and suitable evaluation methodologies, security would be difficult to comprehend, and the cost-benefit trade-offs would be difficult to evaluate. In addition, all the other topic areas can benefit from scalable trustworthy systems, as discussed in Appendix A.

TABLE 1.1: Summary of Gaps, Approaches, and Benefits

Concept	Gaps in Practice	Gaps in R&D	Approaches	Potential Benefits
Requirements	Nonexistent, inconsistent, incomplete nonscalable requirements	Orange Book/Common Criteria have inherent limitations	Canonical, composable, scalable trustworthiness requirements	Relevant developments; Simplified procurement process
System architectures	Inflexibility; Constraints of flawed legacy systems	Evolvable architectures, scalable theory of composability are needed	Scalably composable components and trustworthy architectures	Long-term scalable evolvability maintaining trustworthy operation
Development methodologies and software engineering	Unprincipled systems, unsafe languages, sloppy programming practices	Principles not experientially demonstrated; Good programming language theory widely ignored	Built-in assured scalably composable trustworthiness	Fewer flaws and risks; Simplified evaluations
Analytic tools	Ad-hoc, piecemeal tools with limited usefulness	Tools need sounder bases	Rigorously based composable tools	Eliminating many flaws
Whole-system evaluations	Impossible today for large systems	Top-to-bottom, end-to-end analyses needed	Formal methods, hierarchical staged reasoning	Scalable incremental evaluations
Operational practices	Enormous burdens on administrators	User and administrator usability are often ignored	Dynamic self-diagnosis and self-healing	Simplified, scalable operational management

What are the challenges that must be addressed?

The absence of sound systemwide architectures designed for trustworthiness and the relatively large costs of full verification and validation (V&V) have kept any secure computing base from economically providing the requisite assurance and functionality. (The sole exception is provided by "high-consequence" government applications, in which cost is a secondary concern to national security.) This situation is exacerbated by the scale and complexity often needed to provide required functionality. In addition, the length of the evaluation process can exceed the time available for patches and system upgrades and retarded the incorporation of high assurance information technology. Time-consuming evaluations of trustworthy systems today create long delays when compared with conventional system developments with weaker evaluations. Consequently, development of trustworthy systems can be expected to take longer than is typically planned for COTS systems. In addition, the performance of trustworthy systems typically lags the performance of COTS systems with comparable functions.

One of the most pressing challenges involves designing system architectures that minimize how much of the system must be trustworthy—i.e., minimizing the size and extent of the trusted computing base (TCB). In contrast, for a poorly designed system, any failure could compromise the trustworthiness of the entire system. Designing complex secure systems from the ground up is an exceptionally hard problem, particularly since large systems may have catastrophic flaws in their design and implementation that are not discovered until late in development, or even after deployment. Catastrophic software flaws may occur even in just a few lines of mission-critical code, and are almost inevitable in the tens of millions of lines of code in today's systems. Given the relatively minuscule size of programs and systems that have been extensively verified and the huge size of modern systems and applications, scaling up formal approaches to production and verification of bug-free systems seems like a Herculean task. Yet,

TABLE 1.2: Scalable Trustworthy Systems Overview

Vision: Make the development of trustworthy systems of systems (TSoS) practical; ensure that even very large and complex systems can be built with predictable scalability and demonstrable trustworthiness, using well-understood composable architectures and well-designed, soundly developed, assuredly trustworthy components.

Challenges: Most of today's systems are built out of untrustworthy legacy systems using inadequate architectures, development practices, and tools. We lack appropriate theory, metrics of trustworthiness and scalability, sound composable architectures, synthesis and analysis tools, and trustworthy building blocks.

Goals: Sound foundations and supporting tools that can relate mechanisms to policies, attacks to mechanisms, and systems to requirements, enabling facile development of composable TSoS systematically enhancing trustworthiness (i.e., making them more trustworthy than their weakest components); documented TSoS developments, from specifications to prototypes to deployed systems.

MILESTONES

Incremental Systems	Clean-Slate Systems	Hybrid Systems
Near-term milestones: Sound analytic tools Secure bootloading Trusted platforms	**Near-term milestones:** Alternative architectures Well-specified requirements Sound kernels/VMMs	**Near-term milestones:** Mix-and-match systems Integration tools Evaluation strategies
Medium-term milestones: Systematic use of tools More tool development	**Medium-term milestones:** Provably sound prototypes Proven architectures	**Medium-term milestones:** Use in infrastructures Integration experiments
Long-term milestones: Extensively evaluated systems	**Long-term milestones:** Top-to-bottom formal evaluations	**Long-term milestones:** Seamless integration of COTS/open-source components

Test/evaluation: Identify measures of trustworthiness, composability, and scalability, and apply them to real systems.

Tech transfer: Publish composition methodologies for developing TSoS with mix-and-match components. Release open-source tools for creating, configuring, and maintaining TSoS. Release open-source composable, trustworthy components. Publish successful, well-documented TSoS developments. Develop profitable business models for public-private TSoS development partnerships for critical applications, and pursue them in selected areas.

formally inspired approaches may be more promising than any of the less formal approaches attempted to date. In addition, considerable progress is being made in analyzing system behavior across multiple layers of abstraction. On the other hand, designing complex trustworthy systems and "compromise-resilient" systems on top of insecure components is almost certainly an even harder problem.

As one example, securing the bootload process would be very valuable, but the underlying general principle is that every module of executable software within a system should be backed by a chain of trust, assuring (a) that the integrity of the executable code has not been compromised and (b) that the code resides in memory in a manner that it can be neither read nor altered, but only executed. Firmware residing in ROM, when ROM updating is cryptographically protected for integrity, meets these criteria. Software that is cryptographically protected for integrity, validated

when loaded, and protected by hardware so it can only be executed also meets these criteria.

One of the most relevant challenges for this topic area is how to achieve highly principled system development processes based on detailed and farsighted requirements and sound architectures that can be composed out of demonstrably trustworthy components and subsystems, and subjected to rigorous software, hardware, and system engineering disciplines for its implementation. The tools currently being used do not even ensure that a composed system is at least as trustworthy as its components.

Measuring confidentiality and integrity flaws in trustworthy system construction requires the ability to identify and measure the channels through which information can leak out of a system. Covert channels have been well studied in the constrained, older, local sense of the term. In an increasingly connected world of cross-domain traffic, distributed covert channels become increasingly available. For more distributed forms of covert channels or other out-of-band signaling channels, we lack the science, mathematics, fundamental theory, tools for risk assessment, and the ability to seal off such adverse channels.

Legacy constraints on COTS software, lack of networking support, and serious interoperability constraints have retarded progress. Meaningful security has not been seen as a competitive advantage in the mainstream. Even if trustworthiness were seen in that light, there are no accepted methodologies for design, implementation, operation, and evaluation that adequately characterize the trade-offs among trustworthiness, functionality, cost, and so on.

What approaches might be desirable?

Currently, searching for flaws in microprocessor design makes effective use of formal verification tools to evaluate a chip's logic design, in addition to other forms of testing and simulation. This technology is now becoming very cost-effective. However, it is not likely to scale up by itself to the evaluation of entire hardware/software systems, including their applications. Also, it is unclear whether existing hardware verification tools are robust against nation-state types of adversaries. Formal verification and other analytic tools that can scale will be critical to building systems with significantly higher assurance than today's systems. Better tools are needed for incorporating assurance in the development process and for automating formal verification. These tools may provide the functionality to build a secure computing base to meet many of users' needs for assurance and functionality. They should be available for pervasive use in military systems, as well as to commercial providers of process control systems, real-time operating systems, and application environments. Tools that can scale up to entire systems (such as national-scale infrastructures) will require rethinking how we design, build, analyze, operate, and maintain systems; addressing requirements; system architectures; software engineering; programming and specification languages; and corresponding analysis techniques. System design and analysis, of course, must also anticipate desired operational practice and human usability. It must also encompass the entire system life cycle and consider both environmental adversaries and other adverse influences.

Recent years have seen considerable progress in model checking and theorem proving. In particular, significant progress has been made in the past decade on static and dynamic analysis of source code. This progress needs to be extended, with particular emphasis on realistic scalability that would be applicable to large-scale systems and their applications.

Verification of a poorly built system after the fact has never been accomplished, and is never likely to work. However, because we cannot afford to scrap our existing systems, we must seek an evolutionary strategy that composes new systems out of combinations of old and new subsystems, while minimizing the risks from the old systems. A first step might involve a more formal understanding of the security limitations and deficiencies of important existing components, which would at least allow us to know the risks being taken by using such components in trustworthy composable systems. The ultimate goal is to replace old systems gradually and piecewise over time, to increase trustworthiness for progressively more complex systems.

Verification is expensive. Most COTS systems are built around functionality rather than trustworthiness, and

are optimized on cost of development and time to deployment—generally to the detriment of trustworthiness and often resulting in undetected vulnerabilities. An alternative approach is to start from a specification and check the soundness of the system as it is being built. The success of such an approach would depend on new languages, environments that enable piecewise formal verification, and more scalable proof-generation technology that requires less user input for proof-carrying code. A computer automated secure software engineering environment could greatly facilitate the construction of secure systems. Better yet, it should encompass hardware and total system trustworthiness as well.

Another critical element is the creation of comprehensible models of logic and behavior, with comprehensible interfaces so that developers can maintain an understanding of systems even as they increase in size and scale. Such models and interfaces should help developers avoid situations where catastrophic bugs lurk in the complexity of incomprehensible systems or in the complexity of the interactions among systems. Creation of a language for effectively specifying a policy involving many components is a hard problem. Problems that emerge from interactions between components underscore the need for verifying behavior not only in the lab, but in the field as well.

Finally, efficiently creating provably trustworthy systems will require creation of secure but flexible components, and theories and tools for combining them. Without a secure computing foundation, developers will forever remain stuck in the intractable position of starting from scratch each time. This foundation must include verified and validated hardware, software, compilers, and libraries with easily composable models that include responses to environmental stimuli, misconfigurations and other human errors, and adversarial influences, as well as means of verifying compositions of those components.

What R&D is evolutionary and what is more basic, higher risk, game changing?

Evolutionary R&D might include incremental improvements of large-scale systems for certain critical national infrastructures and specific application domains, such as DNS and DNSSEC, routing and securing the Border Gateway Protocol (BGP), virtualization and hypervisors, network file systems and other dedicated servers, exploitation of multicore architectures, and web environments (e.g., browsers, web servers, and application servers such as WebSphere and WebLogic). However, approaches such as hardening particularly vulnerable components or starkly subsetting functionality are inherently limited, and belief in their effectiveness is full of risks. Goals of this line of R&D include identifying needs, principles, methodologies, tools, and reusable building blocks for scalable trustworthy systems development.

More basic, higher-risk, game-changing R&D broadly includes various topics under the umbrella of composability, because it is believed that only effective composability for trustworthiness can achieve true scalability (just as composability of function enables scalability of system development today). Fundamental research in writing security specifications that are precise enough to be verified, strict enough to be trusted, and flexible enough to be implemented will be crucial to major advances in this area.

Resources

As noted above, this topic is absolutely fundamental to the other topics. The costs of not being able to develop scalable trustworthy systems have already proven to be enormous and will continue to escalate. Unfortunately, the costs of developing high-assurance systems in the past have been considerable. Thus, we must reduce those costs without compromising the effectiveness of the development and evaluation processes and the trustworthiness of the resulting systems. Although it is difficult to assess the costs of developing trustworthy systems in the absence of soundly conceived building blocks, we are concerned here with the costs of the research and prototype developments that would demonstrate the efficacy and scalability of the desired approaches. This may seem to be a rather open-ended challenge. However, incisive approaches that can increase composability, scalability, and trustworthiness are urgently needed, and even relatively small steps forward can have significant benefits.

To this end, many resources will be essential. The most precious resource is undoubtedly the diverse collection of people who could contribute. Also vital are suitable languages for requirements, specification, programming, and so on,

along with suitable development tools. In particular, theories are needed to support analytic tools that can facilitate the prediction of trustworthiness, inclusion modeling, simulation, and formal methods.

Measures of success

Overall, the most important measure of success would be the demonstrable avoidance of the characteristic system failures that have been so common in the past (e.g., see [Neu1995]), just a few of which are noted earlier in this section.

Properties that are important to the designers of systems should be measured in terms of the scale of systems that can be shown to have achieved a specified level of trustworthiness. As noted at the beginning of this section, trustworthiness typically encompasses requirements for security, reliability, survivability, and many other system properties. Each system will need to have its own set of metrics for evaluation of trustworthiness, composability, and scalability. Those metrics should mirror generic requirements, as well as any requirements that are specific to the intended applications. The effectiveness of any computer automated secure software engineering environment (including its generalization to hardware and systems) should be measured in the reduction of person-hours required to construct and verify systems of comparable assurance levels and security. The reuse and size of components being reused should be measured, since the most commonly used components in mission-critical systems should be verified components. Evaluation methodologies need to be developed to systematically exploit the metrics. The measures of success for scalable trustworthy systems also themselves need to be composable into enterprise-level measures of success, along with the measures contained in the sections on the other topic areas that follow.

What needs to be in place for test and evaluation?

Significant improvements are necessary in system architectures, development methodologies, evaluation methodologies, composable subsystems, scalability, and carefully documented, successful worked examples of scalable prototypes. Production of a reasonable number of examples will typically require that will not all succeed. Test and evaluation could proceed for any systems in the context of the exemplars noted above, initially with respect to prototypes and potentially scaling upward to enterprises.

To what extent can we test real systems?

In general, it may be more cost-effective to carry out R&D on components, composability, and scalability in trustworthy environments at the subsystem level than in general system environments. However, composition still requires test and evaluation of the entire system. In that it is clearly undesirable to experiment with critical systems such as power grids, although owners of these systems have realistic but limited-scale test environments. There is considerable need for better analytic tools and testbeds that closely represent reality. Furthermore, if applicable principles, techniques, and system architectures can be demonstrated for less critical systems, successful system developments would give insights and inspiration that would be applicable to the more critical systems without having to test them initially in more difficult environments.

References

[Can2001] Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols (http://eprint.iacr.org/2000/067), 2005. An extended version of the paper from the 42nd Symposium on Foundations of Computer Science (FOCS'01) began a series of papers applying the notion of universal composability to cryptography. Much can be learned from this work regarding the more general problems of system composability.

[Neu1995] Peter G. Neumann. Computer-Related Risks, Addison-Wesley/ACM Press, New York, 1995. See also an annotated index to online sources for the incidents noted here, as well as many more recent cases (http://www.csl.sri.com/neumann/illustrative.html).

[Neu2004] Peter G. Neumann. Principled assuredly trustworthy composable architectures. DARPA-CHATS Final Report, SRI International, Menlo Park, California, December 2004 (http://www.csl.sri.com/neumann/chats4.html). This report characterizes many of the obstacles that must be overcome in achieving composability with predictable results.

[Sal+2009] J.H. Saltzer and F. Kaashoek. Principles of computer design. Morgan Kauffman, 2009. (Chapters 1-6; Chapters 7-11 are online at: http://ocw.mit.edu/ans7870/resources/system/index.htm).

2. Enterprise-Level Metrics (ELMs)

BACKGROUND

What is the problem being addressed?

Defining effective metrics for information security (and for trustworthiness more generally) has proven very difficult, even though there is general agreement that such metrics could allow measurement of progress in security measures and at least rough comparisons between systems for security. Metrics underlie and quantify progress in all other roadmap topic areas. We cannot manage what we cannot measure, as the saying goes. However, general community agreement on meaningful metrics has been hard to achieve, partly because of the rapid evolution of information technology (IT), as well as the shifting locus of adversarial action.

Along with the systems- and component-level metrics that are discussed elsewhere in this document and the technology-specific metrics that are continuing to emerge with new technologies year after year, it is essential to have a macro-level view of security within an organization. A successful research program in metrics should define a security-relevant science of measurement. The goals should be to develop metrics to allow us to answer questions such as the following:

- How secure is my organization?
- Has our security posture improved over the last year?
- To what degree has security improved in response to changing threats and technology?
- How do we compare with our peers with respect to security?
- How secure is this product or software that we are purchasing or deploying?
- How does that product or software fit into the existing systems and networks?
- What is the marginal change in our security (for better or for worse), given the use of a new tool or practice?
- How should we invest our resources to maximize security and minimize risks?
- What combination of requirement specification, up-front architecture, formal modeling, detailed analysis, tool building, code reviews, programmer training, and so on, would be most effective for a given situation?
- How much security is enough, given the current and projected threats?
- How robust are our systems against cyber threats, misconfiguration, environmental effects, and other problems? This question is especially important for critical infrastructures, national security, and many other large-scale computer-related applications.

Enterprise-level metrics (ELMs) address the security posture of an organization and complement the component-level metrics examined elsewhere in the roadmap topics. "Enterprise" is a term that encompasses a wide range. It could in principle apply to the Internet as a whole, but realistically it is intended here to scale in scope from a large corporation or department of the federal government down to the small office/home office (SOHO). For our purposes, an enterprise has a centralized decision making authority to ensure the use of ELMs to rationally select among alternatives to improve the security posture of that enterprise. ELMs can support decisions such as whether adoption of one technology or another might improve enterprise security. ELMs also provide the basis for accurate situational awareness of the enterprise's security posture.

In this discussion, we define metrics relevant to systems and networking within an enterprise, and consider composing host-level and other lower-layer measurements up to an enterprise level. In other words, the goals of ELMs are to understand the security of a large-scale system—enabling us to understand enterprise security as a whole, with a goal of using these measurements to guide rational investments in security. If these ELM goals are met, then extensions to other related cases, such as Internet service providers (ISPs) and their customers, should be feasible.

Security itself is typically poorly defined in real systems, or is merely implicit. One view might be to define it as the probability that a system under attack will meet its specified objectives for a specified period of time in a specified environment. Note that this definition incorporates a specification of system objectives and a specification of the system environment, which would include some notion of a threat model. Although this type of probability metric has been computed for system reliability and for certain system risk assessments, the potential accuracy of such assessments with respect to security seems to be extremely questionable, given the rapidly changing threat environment for IT systems. For example, a presumed high probability of meeting security objectives essentially goes to zero at the instant security exploits are announced and immediately perpetrated.

Security metrics are difficult to develop because they typically try to measure the absence of something negative (e.g., lack of any unknown vulnerabilities in systems and lack of adversary capabilities to exploit both known and unknown vulnerabilities). This task is difficult because there are always unknowns in the system and the landscape is dynamic and adversarial. We need better definitions of the environment and attacker models to guide risk-based determination. These are difficult areas, but progress is achievable.

The following definition from NIST may provide useful insights.

"IT security metrics provide a practical approach to measuring information security. Evaluating security at the system level, IT security metrics are tools that facilitate decision making and accountability through collection, analysis, and reporting of relevant performance data. Based on IT security performance goals and objectives, IT security metrics are quantifiable, feasible to measure, and repeatable. They provide relevant trends over time and are useful in tracking performance and directing resources to initiate performance improvement actions." [http://www.itl.nist.gov/lab/bulletns/bltnaug03.htm]

Most organizations view the answers to the questions listed above in the short term from a financial mind-set and attempt to make cost-benefit trade-off analyses. However, in the absence of good metrics, it is unclear whether those analyses are addressing the right problems. Decisions resulting from such analyses will frequently be detrimental to making significant security improvements in the long term and thus eventually require costly new developments.

What are the potential threats?

Lack of effective ELMs leaves one in the dark about cyberthreats in general. With respect to enterprises as a whole, cybersecurity has been without meaningful measurements and metrics throughout the history of information technology. (Some success has been achieved with specific attributes at the component level.) This lack seriously impedes the ability to make enterprise-wide informed decisions of how to effectively avoid or control innumerable known and unknown threats and risks at every stage of development and operation.

Who are the potential beneficiaries? What are their respective needs?

In short, everyone who is affected by an

automated IT system has the potential to benefit from better security metrics, especially at the enterprise level. Sponsors of security R&D require such metrics to measure progress. With such metrics, decision makers, acquisition managers and investors in security technology could make a better business case for such technology, and guide intelligent investment in such technology. This demand of course would guide the market for development of measurably more secure systems. Metrics can be applied not just to technology, but to practices as well, and can provide management with an incentive structure oriented toward security performance improvement. Robust metrics would enhance the certification and accreditation process, moving toward quantitative rather than qualitative processes. Metrics also can be used to assess the relative security implications of alternative security measures, practices, or policies.

Administrators require metrics to guide the development of optimal network configurations that explicitly consider security, usability, cost, and performance. There seems to be a potential market in insurance and underwriting for predicting and reducing damages caused by cyber attacks, which might be enhanced with the existence of meaningful metrics. However, that market is perhaps undercut not by the lack of suitable metrics, but more by the prevalence of insecure systems and their exploitations and by a historical lack of consistent actuarial data.

Metrics defined relative to a mission threat model are necessary to understand the components of risk, to make risk calculations, and to improve decision making in response to perceived risk. A risk model must incorporate threat information, the value of the enterprise information being protected, potential consequences of system failure, operational practices, and technology. More specifically, risk assessment needs a threat model (encompassing intent and capabilities), a model of actual protective measures, a model of the probability that the adversary will defeat those protective measures, and identification of the consequences of concern or adversary goals. These consequences of concern are typically specific to each enterprise, although many commonalities exist. For critical infrastructures, loss of system availability may be the key concern. For commercial enterprises, loss of proprietary information may be a greater concern than short-term economic losses caused by system outages. Potential beneficiaries, challenges, and needs are summarized in Table 2.1.

What is the current state of the practice?

At present, the practice of measuring security is very ad hoc. Many of the processes for measurement and metric selection are mostly or completely subjective or procedural, as in evaluation of compliance with Sarbanes-Oxley, HIPAA, and so on. New approaches are introduced continually as the old approaches prove to be ineffective. There are measurements such as size and scope of botnets, number of infections in a set of networks, number of break-ins, antivirus detection rates over time, and numbers of warrants served, criminal convictions obtained, and national security letters issued (enforcement). These are not related to fundamental characteristics of systems, but are more about what can be measured about adversaries. Examples include websites that attempt to categorize the current state of the Internet's health, the current state of virus infections world wide, or the number and sizes of botnets currently active.

TABLE 2.1: Beneficiaries, Challenges, and Needs

Beneficiaries	Challenges	Needs
Developers	Establishing meaningful ELMs (comprehensive, feasibly implementable, realistic)	Specification languages, analysis tools for feasibility, hierarchical evaluation, and incremental change
System procurers	Insisting on the use of meaningful ELMs	Certified evaluations
User communities	Having access to the evaluations of meaningful ELMs	Detailed evaluations spanning all relevant aspects of trustworthiness

Numerous initiatives and projects are being undertaken to improve or develop metrics for all or a specific portion of the security domain. Included in these are the following:

- Several government documents and efforts (for example, NIST SP800-55) that describe an approach to defining and implementing IT security metrics. Although some of the measures and metrics are useful, they are not sufficient to answer the security questions identified earlier in this section.
- Methods that assess security based on system complexity (code complexity, number of entry points, etc.). These may give some indication of vulnerability, but in the absence of data on attack success rates or the efficacy of mitigation efforts, these methods prove very little.
- Red Teaming, which provides some measure of adversary work factor and is currently done in security assessments and penetration testing. One can apply penetration testing, using a variety of available tools and/or hiring a number of firms that provide this as a service. For example, this can provide metrics on adversary work factor and residual vulnerabilities before and after implementation of a security plan.
- Heuristic approaches to provide metrics in a number of security-related areas. For example, systems often report a measure of "password strength" (usually on some sort of thermometer). However, password strength is a rather vacuous concept in systems with inherently weak security in other areas.
- Security implementation metrics, which might be used to assess how many systems in an enterprise install a newly announced patch, and how quickly.
- Initiatives in security processes, which might define metrics relating to the adoption of those processes and require extensive documentation. However, such approaches typically are about process and not actual performance improvement with respect to security.

This section focuses on metrics for cybersecurity issues. However, it is also useful to consider existing metrics and design techniques for physical security systems, and the known limitations of those techniques. This information would help advance cybersecurity research. It will also be required as our logical and physical cybersecurity systems become ever more intertwined and interdependent. Similarly, techniques for financial risk management may also be applicable to cybersecurity.

What is the status of current research?

There are initiatives aimed at developing new paradigms for identifying measures and metrics. Some of them attempt to apply tools and techniques from other disciplines; others attempt to approach the problem from new directions. These initiatives include the following:

- **Measures of effectiveness.** The Institute for Defense Analyses (IDA) developed a methodology for determining the effectiveness of cybersecurity controls based on its well-used and -documented methodology for determining the effectiveness of physical security controls. Using a modified Delphi technique, the measures of effectiveness of various components and configurations were determined, which then allowed for a security "ranking" of the potential effectiveness of various architectures and operating modes against different classes of adversaries [IDA2006].
- **Ideal-based metrics.** The Idaho National Laboratory (INL) took a vastly different approach to developing metrics. It chose to specify several best-case outcomes of security and then attempt to develop real-world measures of those "ideals." The resulting set of 10 system measurements covering 7 ideals is being tested in the field to determine how well they can predict actual network or system security performance [McQ2008].
- **Goal-oriented metrics.** Used primarily in the software development domain, the goal-oriented paradigm seeks to establish explicit measurement goals, define sets of questions that relate to achieving the goals, and identify metrics that help to answer those questions.
- **Quality of Protection (QoP).** This is a recent approach that is in early stages of maturity. It

has been the subject of several workshops but is still relatively qualitative [QoP2008].

- **Adversary-based metrics.** MIT Lincoln Laboratory chose to explore the feasibility and effort required for an attacker to break into network components, by examining reachability of those components and vulnerabilities present or hypothesized to be present. It and others have built tools employing attack graphs to model the security of networks.

FUTURE DIRECTIONS

On what categories can we subdivide this topic?

For the purposes of this section, we divide the topic of enterprise-level metrics into five categories: **definition**, **collection**, **analysis**, **composition**, and **adoption**.

Definition

Definition identifies and develops the models and measures to create a set of security primitives (e.g., for confidentiality, integrity, availability, and others). NIST SP 800-55 provides a useful framework for metrics definition. This publication proposes development of metrics along the dimensions of implementation (of a security policy), effectiveness/efficiency, and mission impact.

Ideally, metrics would be defined to quantify security, but such definitions have been difficult to achieve in practice. At the basic level, we would like to quantify the security of systems, answering questions such as the degree to which one system is more secure than another or the degree to which adoption of security technology or practice makes a system more secure. However, as noted above, these measurements are relative to assumed models for adversary capabilities and goals, and to our knowledge of our systems' vulnerabilities—and therefore are potentially limited by shortcomings in the models, requirements, knowledge, assumptions, and other factors.

While this section is focused on enterprise-level metrics (ELMs), we must also consider definitions of metrics for interconnected infrastructure systems, as well as for non-enterprise devices. We must also anticipate the nature of the enterprise of the future; for example, technology trends imply that we should consider smart phones as part of the enterprise. Infrastructure systems may be thought of as a particular class of enterprise-level systems. However, the interrelationships among the different infrastructures also suggest that we must eventually be able to consider meta-enterprises.

Collection

Collection requirements may inspire new research in hardware and software for systems that enable the collection of data through meaningful metrics, ideally in ways that cannot be compromised by adversaries. This includes conditioning the data via normalization, categorization, prioritization, and valuation. It might also include system developments with built-in auditability and embedded forensics support, as well as other topic areas, such as malware defense and situational understanding.

Analysis

Analysis focuses on determining how effectively the metrics describe and predict the performance of the system. The prediction should include both current and postulated adversary capabilities. There has been relatively little work on enterprise-level analyses, because a foundation of credible metrics and foundational approaches for deriving enterprise-level evaluations from more local evaluations have been lacking.

Composition

Since security properties are often best viewed as total-system or enterprise-level emergent properties, research is required in the composability of lower-level metrics (for components and subsystems) to derive higher-level metrics for the entire system. This "composable metrics" issue is a key concern for developing scalable trustworthy systems. In addition, the composability of enterprise-level metrics into meta-enterprise metrics and the composability of the resulting evaluations present challenges for the long-term future.

Adoption

Adoption refers to those activities that transform ELM results into a useful form (such as a measurement paradigm or methodology) that can be broadly used—taking systems, processes, organizational constraints, and human factors into account. Monetary and financial considerations may suggest adoption of metrics such as the number of records in a customer database and a cost per record if those records are disclosed. We may also consider financial metrics retrospectively (the cost of a particular compromise, in terms of direct loss,

reputation, remediation costs, etc.). This retrospection would be useful for system designers and for the insurance underwriting concept mentioned previously.

What are the major research gaps?

In spite of considerable efforts in the past, we do not have any universally agreed-upon methodologies to address the fundamental question of how to quantify system security. At a minimum, an evaluation methodology would support hypothesis testing, benchmarking, and adversary models. Hypothesis testing of various degrees of formality, from simple engagements to formal, well-instrumented experiments, is needed to determine the viability of proposed security measures. Benchmarking is needed to establish a system effectiveness baseline, which permits the progress of the system to be tracked as changes are made and the threat environment evolves. Finally, evaluation must include well-developed adversary models that predict how a specific adversary might act in a given context as systems react to that adversary's intrusions or other exploits.

What are some exemplary problems for R&D on this topic?

The range of requirements for metrics in security is broad. R&D may be focused in any of the following areas:

- Choosing appropriate metrics
- Methods for validating metrics
- Methods for metric computation and collection
- Composition models of metrics to determine enterprise values from subsystem metrics
- Scalability of sets of metrics
- Developing or identifying metric hierarchies
- Measures and metrics for security primitives
- Appropriate uses of metrics (operations, evaluation, risk management, decision making)
- Ability to measure operational security values
- Measuring human-system interaction (HSI)
- Tools to enhance and automate the above areas in large enterprises

What R&D is evolutionary, and what is more basic, higher risk, game changing?

Composability advances (for multiple metrics) could be game-changing advances. Hierarchical composition of metrics should support frameworks such as argument trees and security cases (analogous to safety cases in complex mechanical systems, such as aircraft).

Identifying comprehensive metrics, or a different set of measurement dimensions, might provide a leap forward. The well-known and well-used confidentiality, integrity, availability (CIA) model is good for discussing security, but may not be easily or directly measured in large enterprises. It is also inherently incomplete. For example, it ignores requirements relating to accountability, auditing, real-time monitoring, and other aspects of trustworthiness, such as system survivability under threats that are not addressed, human safety, and so on.

Adapting approaches to metrics from other disciplines is appropriate, but the result is not complete and often not sufficiently applicable (as in the case of probability metrics for component and system reliability). We should consider connections with other fields, while remaining aware that their techniques may not be directly applicable to cybersecurity because of intelligent adversaries and the fluid nature of the attack space.

Many disciplines (such as financial metrics and risk management practices; balanced scorecard, six-sigma, and insurance models; complexity theory; and data mining) operate in environments of decision making under uncertainty, but most have proven methods to determine risk. For example, the field of finance has various metrics that help decision makers understand what is transpiring in their organizations. Such metrics can provide insight into liquidity, asset management, debt management, profitability, and market value of a firm. Capital budgeting tools determining net present-values and internal rates of return allow insights into the returns that can be expected from investments in different projects. In addition, there are decision-making approaches, such as the Capital Asset Pricing Model and options pricing models, that link risk and return to provide a perspective of the entire financial portfolio under a wide range of potential market conditions. These methodologies have demonstrated some usefulness and have been applied across industries to support decision making. A possible analog for

IT security would be sound systems development frameworks that support enterprise-level views of an organization's security. Research is needed to identify system design elements that enable meaningful metrics definition and data collection. Research is also needed on issues in collection logistics, such as the cost of collection and its impact on the metric being used (e.g., whether the collection compromises security).

Research on metrics related to adversary behaviors and capabilities needs to be conducted in several key areas, such as the following:

- The extent of an adversary's opportunity to affect hardware and software needs to be studied. This may lead to research into, for example, global supply-chain metrics that account for potential adversarial influence during acquisition, update, and remote management cycles.
- Metrics in the broad area of adversary work factor have been considered for some time. The simple example is the increase in the recommended length of cryptographic keys as computational power has increased. This work should continue, but there is a question as to the repeatability of the obtained metric.
- Research related to an adversary's propensity to attempt a particular attack, in response to a defensive posture adopted by the enterprise, needs to be conducted.
- Economic or market analysis of adversary actions may provide an indirect metric for security effectiveness. If the cost to exploit a vulnerability on a critical and widely used server system increases significantly, we might surmise that the system is becoming more secure over time or that the system has become more valuable to its adversaries. This approach can be confounded by, for example, the monetary assets accessible to the adversary by compromising the service. (A very secure system not widely used in an attractive target space may discourage a market for high-priced vulnerabilities.) It is also not obvious that this is an enterprise-level metric. Nonetheless, the assembled experts considered market analysis a novel and interesting avenue of research.
- Metrics relating to the impact of cybersecurity recommendations on public- and private-sector enterprise-level systems.

Metrics can guide root-cause analysis in the case of security incidents. Research using existing events should compile a list of metrics that might have avoided the incident if they had been known before the incident.

A stretch objective in the long term is the development of metrics and data collection schemes that can provide actuarial-quality data with respect to security. This is needed for a robust market for insurance against cybersecurity-related risks. Another long-range stretch goal would be to unify the metrics and evaluation methodologies for security of the information domain with the metrics and evaluation methodologies for physical, cognitive, and social domains.

Resources

Industry trends such as exposure to data breaches are leading to the development of tools to measure the effectiveness of system implementations. Industry mandates and government regulations such as the Federal Information Security Management Act (FISMA) and Sarbanes-Oxley require the government and private-sector firms to become accountable in the area of IT security. These factors will lead industry and government to seek solutions for the improvement of security metrics.

Government investment in R&D is still required to address the foundational questions that have been discussed, such as adversary capabilities and threat measurements.

Measures of success

The ability to accurately and confidently predict the security performance of a component, network, or enterprise is the ultimate measure of success for metrics R&D. Interim milestones include better inputs for risk calculation and security investment decisions. The extent to which the evaluation of local metrics (e.g., see the other sections) can be combined into enterprise-level metrics would be a significant measure of success.

What needs to be in place for test and evaluation?

Testbeds and tools within the testbeds are needed to evaluate the descriptive and predictive value and effectiveness of proposed measures and models, particularly for potentially destructive events. Repositories of measurement "baselines" to compare new metric methods and models will also be required. Virtualization and honeynet environments permit assessment of "time to compromise" experimental metrics, possibly considering systems that are identical except for some security enhancement.

Evaluation and experimentation are essential to measure something that is relevant to security. Evaluation methodology goes hand in hand with metrics, and tools that accurately measure and do not distort quantities of interest also have direct influence on metrics.

To what extent can we test real systems?

An enterprise is a testbed of sorts to glean insights on usability, organizational behavior, and response to security practices. Much of the initial collection and verification must be done on real systems to ensure applicability of the measurements and derived metrics.

References

[And2008] R. Anderson. *Security Engineering: A Guide to Building Dependable Distributed Systems*. Wiley, Indianapolis, Indiana, 2008.

[Avi+2004] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1):11-33, January-March 2004.

[Che2006] E. Chew, A. Clay, J. Hash, N. Bartol, and A. Brown. *Guide for Developing Performance Metrics for Information Security*. NIST Special Publication 800-80, National Institute of Standards and Technology, Gaithersburg, Maryland, May 2006.

[CRA2003] Four Grand Challenges in Trustworthy Computing: Second in a Series of Conferences on Grand Research Challenges in Computer Science and Engineering. Computing Research Association, Washington, D.C., 2006 (http://www.cra.org/reports/trustworthy.computing.pdf).

[Jaq2007] A. Jaquith. *Security Metrics*. Addison Wesley Professional, Upper Saddle River, New Jersey, 2007.

[IDA2006] Institute for Defense Analysis. National Comparative Risk Assessment Pilot Project. Draft Final, September 2006, IDA Document D-3309.

[McQ2008] M.A. McQueen, W.F. Boyer, S. McBride, M. Farrar, and Z. Tudor. Measurable control system security through ideal driven technical metrics. *In Proceedings of the SCADA Scientific Security Symposium*, January 2008.

[Met2008] Metricon 3.0, July 29, 2008, with copious URLs (http://www.securitymetrics.org/content/Wiki.jsp?page=Metricon3.0).

[NIS2009] Information Security Training Requirements: A Role- and Performance-Based Model. NIST Special Publication 800-16 Revision 1, National Institute of Standards and Technology, Gaithersburg, Maryland, March 20, 2009 (http://csrc.nist.gov/publications/PubsDrafts.html).

[QoP2008] 4th Workshop on Quality of Protection (Workshop co-located with CCS-2008), October 2008 (http://qop-workshop.org/)

[Swa+2003] M. Swanson, N. Bartol, J. Sabato, J. Hash, and L. Graffo. *Security Metrics Guide for Information Technology Systems*. NIST Special Publication 800-55, National Institute of Standards and Technology, Gaithersburg, Maryland, July 2003.

3. System Evaluation Life Cycle

BACKGROUND

What is the problem being addressed?

The security field lacks methods to systematically and cost-effectively evaluate its products in a timely fashion. Without realistic, precise evaluations, the field cannot gauge its progress toward handling security threats, and system procurement is seriously impeded. Evaluations that take longer than the existence of a particular system version are of minimal use. A suitable **life cycle** methodology would allow us to allocate resources in a more informed manner and enable consistent results across multiple developments and applications.

System evaluation encompasses any testing or evaluation method, including testing environments and tools, deployed to evaluate the ability of a system or a security "artifact" to satisfy its specified critical requirements. A security artifact may be a protocol, device, architecture, or, indeed, an entire system or application environment. Its security depends on the security of the environments in which the artifact will be deployed (e.g., an enterprise or the Internet), and must be reflected throughout the system development life cycle (SDLC). Such a product must meet its specification with respect to a security policy that it is supposed to enforce, and not be vulnerable to attack or exploitation that causes it to perform incorrectly or maliciously. Secondary but also important performance goals can be expressed as "do no harm." The proposed artifact should not inflict collateral damage on legitimate actors or traffic in the Internet, and it should not create additional security problems. The system evaluation life cycle thus denotes continuous evaluation throughout the system life cycle (requirements, design, development and implementation, testing, deployment and operations, and decommissioning and disposal). See [NIS2008].

Security evaluation in the SDLC involves four major areas in addressing potential threats:

- Developing explicit requirements and specifications for systems, including security features, processes, and performance requirements for each development phase in sufficient detail.
- Understanding whether a product meets its specification with respect to a security policy that it is suppose to enforce. A part of this is understanding how well the product meets the specification and ensures that there are no exploitable flaws. In the case of systems enforcing mandatory confidentiality or integrity policies, this includes demonstration of the limits to adversarial exploitation of covert channels.
- Understanding whether a product can be successfully attacked or bypassed by testing it in each phase of its development life cycle, either in a testbed or through a mathematical model or simulation.

- Developing system evaluation processes whereby incremental changes can be tracked and rapidly reevaluated without having to repeat the entire process.

In each case, independent assessment of a product could reduce reliance on vendor claims that might mask serious problems. On the other hand, embedded self-assurance techniques (such as proof-carrying code) could also be used to demonstrate that certain properties were satisfied.

Systematic, realistic, easy-to-use and standardized evaluation methods are needed to objectively quantify performance of any security artifacts and the security of environments where these artifacts are to be deployed, before and after deployment, as well as the performance of proposed solutions. The evaluation techniques should objectively quantify security posture throughout the critical system life cycle. This evaluation should support research, development, and operational decisions, and maximize the impact of the investment.

Finally, evaluation must occur in a realistic environment. The research community lacks data about realistic adversarial behavior, including the tactics and techniques that adversaries use to disrupt and deny normal operations, as well as normal system use patterns and business relationships, to create a realistic environment for evaluation that resembles current environments in which systems are deployed. We also lack understanding of human behavior as users interact with the system and with security artifacts. Such understanding is needed to evaluate the likelihood of human acceptance of proposed security artifacts and to simulate human actions during evaluation (e.g., browsing patterns during evaluation of a web server defense).

What are the potential threats?

Threats against information and information systems are at the heart of the need for robust system evaluation. In addition to the threats to operational systems, adversaries have the potential to affect the security of artifacts at numerous points within the development life cycle. The complexity of systems, modifications, constant changes to supply chains, remote upgrades and patches, and other factors give rise to numerous new threat vectors.

Who are the potential beneficiaries? What are their respective needs?

With regard to the system life cycle, system architects, engineers, developers, and evaluators will benefit from enhanced methods of evaluation. Beneficiaries of improved security evaluations range from large and small enterprises to end users of systems. Although beneficiaries' needs are generally the same—to prevent security incidents and to respond quickly to those that evade prevention and minimize damage, while protecting privacy—environments that they seek to protect may be very different, as are their needs for reliability, correctness of operation, and confidentiality. Direct beneficiaries of better evaluation methods are system developers, system users and administrators; the customers of security products (because they need reliable means to evaluate what they buy); the creators of these products, such as software and hardware companies; and researchers (because they need to measure their success). Having effective evaluation methods opens the door to the possibility of standardization of security and to formation of attestation agencies that independently evaluate and rank security products. The potential beneficiaries, challenges, and needs are summarized in Table 3.1.

What is the current state of the practice?

Evaluation of security artifacts is ad hoc. Current methodologies, such as these discussed in NIST SP800-64 (*Security Considerations in the System Development Life Cycle*) [NIS2008] and Microsoft's *The Security Development Life cycle* [How+2006], merely reorder or reemphasize many of the tools and methods that have been unsuccessful in creating security development paradigms. There are neither standards nor metrics for security evaluation. Product developers and vendors evaluate their merchandise in-house, before release, via different tests that are not disclosed to the public. Often, real evaluation takes place in customer environments by product vendors collecting periodic statistics about threats detected and prevented during live operation. Although this is the ultimate measure of success—how a product performs in the real world—it does not offer security guarantees to customers prior to purchase. There have been many incidents when known security devices have failed (e.g., the Witty worm infected security products from a well-known security product vendor). In

TABLE 3.1: Beneficiaries, Challenges, and Needs

Beneficiaries	Challenges	Needs
System developers	Integrate components into systems with predictable and dependable security properties; effectively track changes from one version to another.	Robust methods to compare components to be used in new systems. Tools, techniques, and standards of system evaluation to enable certification of security properties of products developed.
System owners and administrators	Understand the risk to their information operations and assets. Operate and maintain information systems in a secure manner.	Suites of tools that can be used throughout the operational phases of the system life cycle to evaluate the current system state and the requirements and impacts of system upgrades or changes.
End users	Operate confidently in cyberspace.	Recognized and implemented life cycle system evaluation methods that provide high confidence in the safety and security of using online tools and environments.

addition, past efforts such as evaluations of the Trusted System Security Evaluation Criteria and the Common Criteria [ISO1999] suffer from inadequate incremental methods to rapidly reevaluate new versions of systems.

What is the status of current research?

Relatively little research has been done on system evaluation methods. The research community still values such topics much less than research on novel defenses and attacks. The metrics and measures needed to describe security properties during the evaluation life cycle must be developed. (See Section 2). The lack of metrics results in security products that cannot be compared and in solving past problems instead of anticipating and preventing future threats. Because the necessary metrics are likely to depend on the nature of the threat a security artifact aims to address, it is likely that the set of metrics will be large and complex.

FUTURE DIRECTIONS

On what categories can we subdivide this topic?

We initially discuss this topic relative to a nominal life cycle model. The SDLC phases represented in our nominal model are: requirements, design, development and implementation, testing, deployment and operations, and decommissioning. System evaluation has to be done throughout the entire life cycle, with continuous feedback and reevaluation against previous stages.

Potential R&D directions that might be pursued at multiple life cycle phases include the following:

- Develop cost-effective methods to specify security features for succeeding life cycle phases.
- Develop adversarial assessment techniques that identify and test for abnormal or unintended operating conditions that may cause exploitable vulnerabilities.
- Develop realistic traffic, adversary, and environment models that span all four domains of conflict (physical, information, cognitive, and social).
- Develop security test cases, procedures, and models to evaluate the artifact in each life cycle phase.
- More effectively perform speedy reevaluations of successive versions resulting from changes in requirements, designs, implementation, and experience gained from uses of system applications.

The following discussion considers the individual phases.

Requirements

- Establish a sounder basis for how security requirements get specified, evaluated, and updated at each phase in the life cycle.

- Incorporate relevant (current and anticipated) threats models in the requirements phase so that the final specification can be evaluated against those threats.
- Specify what constitutes secure operation of systems and environments.
- Establish requirement specification languages that express security properties, so that automated code analysis can be used to extract what the code means to do and what its assumptions are.

Design

- Be able to share data with adequate privacy, including data on attacks, and with emphasis on economics of data sharing.
- Develop a richer process to develop data used to validate security claims.
- Develop frameworks for threat prediction based on data about current attacks and trends.
- Develop simulations of (unusual or unanticipated) system states that are critical for security, as opposed to simulation of steady states.

Development and Implementation

- Pursue evaluation methods able to verify that an implementation follows requirements precisely and does not introduce anything not intended by requirements. If specifications exist, this can be done in two steps: verifying consistency of specifications with requirements and then consistency of software with specifications. Concerns about insider threats inside the development process also need to be addressed.
- Pursue verification that a system is implemented in a way that security claims can be tested.
- Consider new programming languages, constraints on or subsets of existing languages, and hardware design techniques that express security properties, enforce mandatory access controls, and specify interfaces, so that automated code analysis can be used to extract what the code means to do and what its assumptions are.

Testing

- Select and evaluate metrics for evaluation of trustworthiness requirements.
- Select and use evaluation methods that are well suited to the anticipated ranges of threats and operational environments.
- Develop automated techniques for identifying all accessible system interfaces (intentional, unintentional, and adversary-induced) and system dependencies. For example, exploitation of a buffer overflow might be considered a simple example of an unintended system interface.
- Develop and apply automated tools for testing all system dependencies under a wide range of conditions. As an example, some adversaries may exploit hardware-software interactions that are ill-documented, are time-dependent, and occur only when all of the subsystems have been integrated.
- Conduct Red Team exercises in a structured way on testbeds to bring realism. Expand the Red Team concept to include all phases of the life cycle.
- Establish evolvable testbeds that are easily upgradeable as technology, threat, and adversary models change.
- Improve techniques for combined performance, usability, and security testing. This includes abnormal environments (e.g., extreme temperatures) and operating conditions (e.g., misuse by insiders) that are relevant for security testing but may exceed the system's intended range of operation.

Deployment and Operations

- Establish and use evaluation methods that can compare actual operational measurements with design specifications to provide feedback to all life cycle phases.
- Develop methods to identify system, threat, or environment changes that require reevaluation to validate compliance with evolving security requirements.
- Define and consistently deploy certification and accreditation methods that provide realistic values regarding the trustworthiness of a system with respect to its given requirements.

Decommissioning

- Develop end-of-life evaluation

methods to verify that security requirements have been achieved during the entire life cycle. This includes ensuring that an adversary can not extract useful information or design knowledge from a decommissioned or discarded security artifact.

- Inform threat models from product or system end-of-life analysis.

What are the major research gaps?

A major gap is lack of the knowledge and understanding of the threat domain that is needed to develop realistic security requirements. One reason for this gap is the lack of widely available data on legitimate and attack traffic, for various threats and at various levels. Another large challenge is the lack of reliable methods to measure success of various attacks, and inversely to measure the success of defensive actions against attacks.

Yet another challenge lies in not understanding how much realism matters for testing and evaluation. For example, can tests in a 100-node topology with realistic traffic predict behavior in a 10,000-node topology, and for which threats? Some large "hybrid" testbeds may need mixtures of real, emulated, and simulated entities to provide flexible tradeoffs between test accuracy and testbed cost/scalability. If so, then workload estimation and workload partitioning tools are needed to design experiments for large testbeds. (A simple example is that a malware research testbed typically needs real hosts but can emulate or simulate the network interconnections.) Also relevant here is the DETERlab testbed (cyber-DEfense Technology Experimental Research laboratory testbed (http://www.isi.edu/deter). The DETERlab testbed is a general-purpose experimental infrastructure for use in research (http://www.deterlab.net).

Understanding of which evaluation methods work for which threats is also lacking. For example, formal reasoning and model checking may work for software, but simulation may work better for routing threats. Finally, there is no peer review mechanism to review and validate evaluation mechanisms or proposals.

What are some exemplary problems for R&D on this topic?

Possible directions to solve current problems in security evaluation are: (a) system architectures that enhance evaluation throughout the development cycle; (b) development of security metrics and benchmarks for components, subsystems, and entire enterprises; (c) development of tools for easy replication of realistic environments in testbeds and simulations; (d) realistic adversary models, including how those adversaries might react to changes in the defensive security posture; and (e) the encompassing methodologies that bring these components together.

Projects envisioned in this area include the following:

- Develop cost-effective methodologies and supporting tools that can result in timely evaluations and can rapidly track the effects of incremental changes.
- Enable creation of attack data repositories under government management, similar to the PREDICT repository (http://www.predict.org), for legitimate data. Develop approaches to bring realism into simulations and testbeds.
- Develop research about when scalability matters, and in what way. Develop research about when realism (or simulation) matters, and what type of realism. Develop research about what type of testing works for which threats and environments. Develop simple metrics for system and network health and for attack success.
- Develop detailed metrics for system and network health and for attack success. Develop benchmarks and standardize testing.

What R&D is evolutionary, and what is more basic, higher risk, game changing?

The development over time of system evaluation tools, methodologies, measures, and metrics will require iterations and refinements of the successes of short-term projects, as well as long-term research. There are short- and long-term implications in many of the projects and challenges noted.

Evolutionary, relatively short-term R&D challenges include the following:

- Defining verifiable parametric sets of requirements for trustworthiness and improved models for assessing requirements.
- Devising methods to recreate realism in testbeds and simulations while providing flexible trade-offs between cost, scalability, and accuracy. (These include better methods for designing experiments for large testbeds).
- Developing methods and representations such as abstraction models to describe threats, so that designers can develop detailed specifications.
- Developing user interfaces, tools, and capabilities to allow complex evaluations to be conducted.
- Developing tool sets that can grow with technology (e.g., 64-bit words, IPv6).
- Creating better techniques for testing combined performance, usability, and security.
- Developing understanding of how much realism matters and what type of realism is possible and useful.

Long-term, high-risk R&D challenges include the following:

- Developing models of correct operation for various network elements and networks at and across all levels of protocol models.
- Developing metrics for attack success and for security based on the models of correct operation.
- Developing benchmarks to standardize testing.
- Developing understanding about advantages and limitations of various evaluation methods (simulation, emulation, pilot deployment, model checking, etc.) when related to specific threats.
- Managing risky test environments (such as those containing malware).
- Developing better techniques for security testing across all domains of conflict.
- Developing integrated, cost-effective methodologies and tools that systemically address all of the above desiderata, including facilitation of scalable trustworthiness (Section 1), survivability (Section 7), resistance to tampering and other forms of insider misuse by developers (Section 4), rapid reevaluation after incremental changes, and suitable uses of formal methods where most usefully applicable—among other needs. The potential utility of formal methods has increased significantly in the past four decades and needs to be considered whenever it can be demonstrably effective.

Resources

Academia and industry should collaborate to share data about traffic, attacks, and network environments and to jointly define standards and metrics for evaluation, including joint design of realism criteria for evaluation environments.

Government should help in mandating, regulating, and promoting this collaboration, especially with regard to data sharing. Legal barriers to data sharing must be addressed. Some industry sectors may be reluctant to share vulnerability data because of legal liability concerns. There may also be privacy and customer relations concerns. An example would be data sharing by common carriers where the shared data uniquely identify individual customers. The government should also provide more complete threat and adversary capability models for use in developing evaluation and testing criteria.

Other potential government activities include the following:

- Propose evaluation methods that are proven correct as national or international standards for tech transfer. They also should be implemented in current popular simulations and testbeds. Industry should be encouraged to use these methods, perhaps via market incentives.
- Form attestation agencies that would evaluate products on the market, using evaluation methods that are ready for tech transfer, and rank those products publicly.
- Create a National CyberSecurity and Safety Board that would collect attack reports from organizations and share them in a privacy-safe manner. The board could also mandate sharing. Another way is establishing

a PREDICT-like repository for attack data sharing. Yet a third way is developing market incentives for data sharing.

- Fund joint academic/industry partnerships in a novel way. Academics have a hard time finding industry partners that are willing to commit to tech transfer. A novel way would have government find several partners from various fields: enterprises, ISPs, government networks, SCADA facilities, security device manufacturers, etc. These partners would pledge to provide data to researchers in the evaluation area and to provide small pilot deployments of technologies developed by projects in other areas (e.g., solutions for critical system availability). Thus, various evaluation methods could be compared with real-deployment evaluations. Without this ground truth comparison, it is impossible to develop good evaluation methods because evaluation must correctly predict ground truth.

Measures of success

One key milestone as a measure of success will be the eventual adoption by standards bodies such as NIST or ISO of consistent frameworks, methodologies, and tools for system evaluation. System developers will be able to choose components from vendors based on results obtained from well-known and well-established evaluation methods. Direct comparisons of vendor products will be possible, based on measures of performance in standard tests.

What needs to be in place for test and evaluation?

A flexible, scalable, and secure large-scale testbed would enable high-fidelity tests of products using new development and evaluation methods.

To what extent can we test real systems?

Because system evaluation must occur at all phases of the life cycle, there should be opportunities to test new tools and methodologies on real systems inobtrusively.

References

[Ade2008] S. Adee. The hunt for the kill switch. IEEE Spectrum, 45(5):32-37, May 2008 (http://www.spectrum.ieee.org/may08/6171).

[DSB2005] *Defense Science Board Task Force on High Performance Microchip Supply*, February 2005 (http://www.acq.osd.mil/dsb/reports/2005-02-HPMS_Report_Final.pdf).

[How+2006] M. Howard and S. Lipner. *The Security Development Life Cycle.* Microsoft Press, Redmond, Washington, 2006.

[ISO1999] International Organization for Standardization / International Electrotechnical Commission (ISO/IEC) International Standard 15408:1999 (parts 1 through 3), *Common Criteria for Information Technology Security Evaluation*, August 1999.

[NIS2008] *Security Considerations in the System Development Lhife Cycle.* NIST Special Publication 800-64 Revision 2 (Draft), National Institute of Standards and Technology, Gaithersburg, Maryland, March 2008.

4. Combatting Insider Threats

BACKGROUND

What is the problem being addressed?

Cybersecurity measures are often focused on threats from outside an organization, rather than threats posed by untrustworthy individuals inside an organization. Experience has shown that insiders pose significant threats:

- Trusted insiders are among the primary sources of many losses in the commercial banking industry.
- Well-publicized intelligence community moles, such as Aldrich Ames, Robert Hanssen, and Jonathan Pollard, have caused enormous and irreparable harm to national interests.
- Many insiders involved in misuses were hired as system administrators, became executives, or held other kinds of privileges [Cap2008.1, Cap2008.2].

This section focuses on insider threats to cyber systems and presents a roadmap for high-impact research that could aggressively curtail some aspects of this problem. At a high level, opportunities exist to mitigate insider threats through aggressive profiling and monitoring of users of critical systems, "fishbowling" suspects, "chaffing" data and services users who are not entitled to access, and finally "quarantining" confirmed malevolent actors to contain damage and leaks while collecting actionable counter-intelligence and legally acceptable evidence.

There are many proposed definitions of the insider threat. For the purposes of this discussion, an **insider threat** is one that is attributable to individuals who abuse granted privileges. The scope of consideration here includes individuals masquerading as other individuals, traitors abusing their own privileges, and innocents fooled by malevolent entities into taking adverse actions. Inadvertent and intentional misuse by privileged users are both within the scope of the definition. Although an insider can have software and hardware acting on his or her behalf, it is the individual's actions that are of primary concern here. Software proxies and other forms of malevolent software or hardware—that is, electronic insiders—are considered in Section 5 on combatting malware and botnets.

The insider threat is context dependent in time and space. It is potentially relevant at each layer of abstraction. For example, a user may be a physical insider or a logical insider, or both. The threat model must be policy driven, in that no one description will fit all situations.

Unlike unauthorized outsiders and insiders who must overcome security controls to access system resources, authorized insiders have legitimate and (depending on their positions) minimally constrained access to computing resources. In addition, highly

trusted insiders who design, maintain, or manage critical information systems are of particular concern because they possess the skills and access necessary to engage in serious abuse or harm. Typical trusted insiders are system administrators, system programmers, and security administrators, although ordinary users may have or acquire those privileges (sometimes as a result of design flaws and implementation bugs). Thus, there are different categories of insiders.

What are the potential threats?

The insider threat is often discussed in terms of threats to confidentiality and privacy (such as data exfiltration). However, other trustworthiness requirements, such as integrity, availability, and accountability, can also be compromised by insiders. The threats span the entire system life cycle, including not only design and development but also operation and decommissioning (e.g., where a new owner or discoverer can implicitly become a de facto insider).

Who are the potential beneficiaries? What are their respective needs?

The beneficiaries of this research range from the national security bodies operating the most sensitive classified systems to homeland security officials who need to share Sensitive But Unclassified (SBU) information/Controlled Unclassified Information (CUI), and to health care, finance, and many other sectors where sensitive and valuable information is managed. In many systems, such as those operating critical infrastructures [Noo+2008], integrity, availability, and total system survivability are of highest priority and can be compromised by insiders.

Beneficiary needs may include tools and techniques to prevent and detect malicious insider activity throughout the entire system life cycle, approaches to minimize the negative impact of malicious insider actions, education and training for safe computing technology and human peer detection of insider abuses, and systems that are resilient and can effectively remediate detected insider exploits. Of particular interest will be the ability to deal with multiple colluding insiders—including detecting potential abuses and responding to them.

What is the current state of the practice?

The insider threat today is addressed mostly with procedures such as awareness training, background checks, good labor practices, identity management and user authentication, limited audits and network monitoring, two-person controls, application-level profiling and monitoring, and general access controls. However, these procedures are not consistently and stringently applied because of high cost, low motivation, and limited effectiveness. For example, large-scale identity management can accomplish a degree of nonrepudiation and deterrence but does not actually prevent an insider from abusing granted privileges.

Technical access controls can be applied to reduce the insider threat but not eliminate it. The technologies traditionally brought forward by the research community are multilevel security (MLS), an example of mandatory access controls (MAC) that prevents highly sensitive information from being accessed by less privileged users. Some work has also been done on multilevel integrity (MLI [Bib1977]), which prevents less trusted entities from affecting more trusted entities. However, these are typically too cumbersome to be usable in all but the most extreme environments; even in such environments, the necessary systems are not readily available. Access controls that are used in typical business environments tend to be discretionary, meaning that the individual or group of individuals who are designated as owners of an object can arbitrarily grant or deny others access to the object. Discretionary access controls (DAC) typically do not prevent anyone with read access to an object from copying it and sharing the copy outside the reach of that user's access control system. They also do not ensure sufficient protection for system and data integrity. Further background on these and other security-related issues can be found in [And08,Bis02,Pfl03].

File and disk encryption may have some relevance to the insider threat, to the extent that privileged insiders might not be able to access the encrypted data of other privileged users. Also of possible relevance might be secret splitting, k-out-of-n authorizations, and possibly zero-knowledge proofs. However, these would need considerable improvement if they were to be effective in commercial products.

What is the status of current research?

Several studies of the insider threat have been produced in the past 10 to 15 years, although some of these rely on research on access controls dating back as many as 40 years. These need to be compiled and serve as input to a taxonomy of the threats and possible violations. Ongoing and emerging research efforts include the following:

- The 2008 Dagstuhl summer seminar on Countering Insider Threats [Dag08] included position papers that are being considered for publication as a book. It represented a broad assessment of a wide variety of considerations.
- Ongoing insider and identity management projects under the aegis of The Institute for Information Infrastructure Protection (I3P)—for example, decoy networking and honeypots, correlating host and network indicators of insider threats, and behavior-based access control. Three papers from the I3P identity management projects were presented at IDtrust 2009. See the references in Section 6.
- Two Carnegie Mellon University reports on insider threats in government [Cap2008.1] and in information technology generally, with emphasis on the financial sector [Cap2008.2]; see also [Ran2004] and [FSS2008].

In addition, a DoD **Insider Threat to Information Systems** report [IAT2008], a study of best practices [HDJ2006], various Columbia University papers and a book on insider threats (e.g., [Sto+2008]), and an NSA/ARDA (IARPA) report on classifications and insider threats [Bra2004] are relevant. Also, the Schonlau data set for user command modeling may be of interest (www.schonlau.net).

FUTURE DIRECTIONS

On what categories can we subdivide this topic?

Approaches for coping with insider misuse can be categorized as collect and analyze (monitoring), detect (provide incentives and data), deter (prevention should be an important goal), protect (maintain operations and economics), predict (anticipate threats and attacks), and react (reduce opportunity, capability, and motivation and morale for the insider). For present purposes, these six categories are grouped pairwise into three bins: collect and analyze, detect; deter, protect; and predict, react.

What are the major research gaps?

Many gaps relating to insider threats need to be better understood and remediated.

- **Checking.** Better mechanisms are needed for policy specification and automated checking (e.g., role-based access control [RBAC] and other techniques). However, any such mechanism must have precise and sound semantics if it is to be useful. (Some past work on digital rights management may be of some indirect interest here.)
- **Response strategy and privacy protection for falsely accused insider abuses.** In particular, privacy-enhanced sharing of behavior models and advanced fishbowling techniques to enable detailed monitoring and limit damage by a suspected inside threat. (See Section 10.)
- **Behavior-based access control.**
- **Decoys, deception, tripwires in the open.**
- **Beacons in decoy (and real) documents.** Adobe and other modern platforms perform a great deal of network activity at startup and during document opening, potentially enabling significant beaconing.
- More **pervasive monitoring** and **profiling**, coupled with remediation in the presence of detected potential misuses.
- **Controlled watermarking** of documents and services to trace sources.
- **Useful data.** The research community needs much more data and more realistic data sets for experimentation.
- **Procedures and technology for emergency overrides** are needed in almost every imaginable application, but must typically be specific to each application. They are particularly important in health care, military, and other situations where human lives depend on urgent access. The existing limitations are in part related to lack of motivation for developing and using fine-grained

access controls. In addition, emergency overrides can be abused by insiders who feign or exploit crises. Overall, approaches must be closely connected to policy specifications.

- **Lessons may be learned from safety systems.** For example, in process control applications, separate safety systems are used to ensure that a process is safely shut down when certain parameters are exceeded because of failure of the control system or for any unanticipated reasons. Analogous protection mechanisms for an information system might ensure that certain operations are never allowed, regardless of the privileges of the users attempting them. Similarly, the principles of least common mechanism and least privilege should be applied more consistently. Also relevant would be "safe booting" for self-protected monitors.
- From a user perspective, **security** and **usability** must generally be integrally aligned, but especially with respect to insider misuse. For example, users should not feel threatened unless they are actually threats to system integrity and to other users. (Interactions with the usability topic in Section 11 are particularly relevant here.)
- **Privacy** is an important consideration, although it typically depends on the specific policies of each organization.
- **Existing access controls** tend to be inadequately fine-grained with respect to preventing insider misuse. In addition, even the existing controls are not used to their full extent. Moreover, better mechanisms are needed for both active monitoring (for detection and response) and passive monitoring (for later analysis and forensics). Note that the prevention/monitoring/recording/archiving mechanisms must themselves be able to withstand threats, especially when the defenders are also the attackers. Also, collection of evidence that will stand up in court is an important part of deterrence. To this end, forensic mechanisms and information must be separated from the systems themselves.

Advanced fine-grained differential access controls; role-based access controls; serious observance of separation of roles, duties, and functionality; and the principle of least privilege also need to be integrated with functional cryptography techniques, such as identity-based and attribute-based encryption, and with fine-grained policies for the use of all the above concepts.

What are some exemplary problems for R&D on this topic?

The categories noted above and some potential approaches are summarized in Table 4.1.

Collect and Analyze

- Data sets relating to insider behavior and insider misuse need to be established. Very few such data sets on insider behavior are available today, in part because victims are reluctant to divulge details and in part because many cases remain unknown beyond local confines. What data should be collected and how it should be made available (openly or otherwise), perhaps via trustworthy third parties, need to be considered. Privacy concerns must be addressed.
- Systems need to be designed to be auditable in ways sufficient to allow collection and analysis of forensic-quality data.
- Models are needed to represent both normal and abnormal insider activity. However, past experience with pitfalls of such models needs to be respected.
- Methodologies are needed for measuring and comparing techniques and tools meant to handle insider threats.

Detect

- Detection of insider abuse and suspected anomalies must be timely and reliable.
- Data mining, modeling, and profiling techniques are needed for detection of malicious insider activity.
- Better techniques are needed to determine user intent from strict observation, as opposed to merely detecting deviations from expected policies.
- Prediction and detection need to be effectively integrated.

TABLE 4.1: Potential Approaches to Combatting Insider Threats

Category	Definition	Potential Approaches
Collect and Analyze, Detect	Understanding and identifying threats and potential risks	Broad-based misuse detection oriented to insiders
Deter, Protect	Trustworthy systems with specific policies to hinder insider misuse	Inherently secure systems with differential access controls
Predict, React	Remediation when insider misuse is detected but not prevented	Intelligent interpretation of likely consequences and risks

Deter

- Fine-grained access controls and correspondingly detailed accountability need to have adequate assurance. Audit logs must be reduced to be correctly interpreted, without themselves leaking information.
- Deterrence policies need to be explored and improved. Training should include use of decoys.
- Incentives need to be developed, such as increased risks of being caught, greater consequences if caught, lessened payoffs if successful, and decreased opportunities for user disgruntlement. The role of an ombudsperson should also be considered in this context.
- Increased incentives for anonymous whistle-blowing, engendering an atmosphere of peer-level misuse detection and monitoring.
- Social, ethical, and legal issues, as well as human factors, need to be addressed in a multidisciplinary fashion.

Protect

- Using a life cycle view could be helpful to establish security perimeters for specific purposes and particular policies, and to identify all relevant insiders therein. (See the section on System Evaluation Life Cycle.) Note that in many cases there are no specific boundaries between inside and outside.
- Continuous user authentication and reauthentication may be desirable to address insider masquerading.
- System architectures need to pervasively enforce the principle of least privilege, which is particularly relevant against insider threats. The principle of least common mechanism could also be useful, restricting functionality and limiting potential damage. Access control mechanisms must move beyond the concept of too-powerful superuser mechanisms, by splitting up the privileges as was done in Trusted Xenix. Mechanisms such as k-out-of-n authorizations might also be useful. New access control mechanisms that permit some of the discipline of multilevel security might also help.
- Deception, diversity, and making certain protection mechanisms more invisible might be useful in addressing the insider threat. Decoys must be conspicuous, believable, differentiable (by good guys), noninterfering, and dynamically changing.
- New research is especially needed in countering multiple colluding insiders. For example, the development of defensive mechanisms that systematically necessitate multiple colluders would be a considerable improvement.
- Anti-tamper technologies are needed for situations where insiders have physical access to systems. Similar technologies may be desirable for logical insiders. Inspiration from nuclear safety controls can illuminate some of the concerns.
- Protections are needed for both system integrity and data integrity, perhaps with finer-grained controls than for outsiders. In addition, operational auditing and rollback mechanisms are needed subsequent to integrity violations. Note that physical means (e.g.,write-once media) and logical means (log-structured file systems) are both relevant.

- Mechanisms are needed that exhaustively describe and enforce the privileges that a user is actually granted. In particular, visualization tools are needed for understanding the implications of both requested and granted privileges, relative to each user and each object. This approach needs to include not just logical privileges, but also physical privileges.
- Mechanisms are needed to prevent overescalation of privileges on a systemwide basis (e.g., chained access that allows unintended access to a sensitive piece of data). However, note that neither trust nor delegation is a transitive operation.

Predict

- Various predictive models are needed—for example, for indicators of risks of insider misuse, dynamic precursor indicators for such misuse, and determining what is operationally relevant (such as the potentially likely outcomes).
- Dynamic analysis techniques are needed to predict a system component's susceptibility to a certain insider attack, based on system operations and configuration changes.
- Profiles of expected good behavior and profiles of possible bad behavior are generally both useful, but neither approach is sufficient. Additional approaches are needed.
- Better technologies are needed to achieve meaningful prediction, including analysis of communications, user behavior, and content. Prediction must address users and surrogates, as well as their actions and targets.

React

- Automated mechanisms are needed that can intercede when misuse is suspected, without jeopardizing system missions and without interfering with other users. For example, some sort of graceful degradation or system recovery may be needed, either before misuse has been correctly identified or afterwards.
- Mechanisms and policies are needed to react appropriately to the detection of potentially actively colluding insiders.
- Architecturally integrated defense and response strategies might mitigate the effects of insider attacks—for example, an insider who is able to override existing policies. One strategy of considerable interest would be unalterable (e.g., once-writable) and non-bypassable audit trails that cannot be compromised. Another strategy would be mechanisms that cannot be altered without physical access, such as overriding safety interlocks.
- Architecturally integrated response strategies might also be invoked when misuse is detected, gathering forensics-worthy evidence of the potential network of inside threats, adversary sources and methods, to enable law-enforcement use of evidence.
- Research is needed on scalable mechanisms for revocable credentials, perfect forward secrecy built into systems, and other approaches that could simplify timely reactions.
- Note that these categories are somewhat interrelated. Any research program related to coping with the insider threats needs to keep this in mind. Table 4.2 summarizes some of the research gaps, research initiatives, benefits, and time-frame.

What are the near-term, mid-term, long-term capabilities that need to be developed?

Near Term

- Compile and compare existing studies relating to the insider threat. (Detect)
- Develop data collection mechanisms and collect data. (Detect)
- Evaluate suitability of existing RBAC R&D to address insider threats. (Protect)
- Develop anti-tampering approaches. (Protect)
- Explore the possible relevance of digital rights management (DRM) approaches. (Protect)

Medium Term

- Develop feature extraction and machine learning mechanisms to find outliers. (Detect)
- Develop tools to exhaustively and accurately understand granted privileges as roles and system configurations change. (Detect)
- Develop procedures to evaluate insider threat protection methods in reliable and comparable ways. (Detect)

TABLE 4.2: Gaps and Research Initiatives

Identified Gap	Research Initiatives	Benefit	Time Frame
Inadequately fine-grained access controls	Better mechanisms, policies, monitoring	Better detection and prevention of insider misuse	Near- to long term
Absence of insider-misuse aware detection	Better detection tools	More precise detection of insider misuse	Near term
Difficulties in remediation	Mixed strategies for finer-grained, continuous monitoring and action	Flexible response to detected misuses	Longer term

- Develop better methods to combat insiders acting alone. (Protect)
- Pursue the relevance and effectiveness of deception techniques. (Protect)
- Incorporate integrity protection into authorization and system architectures. (Protect)
- Develop behavior-based security, for example, advanced decoy networking. (Protect)
- Develop and apply various risk indicators. (React)

Long Term

- Establish effective methods to apply the principle of least privilege. (Protect)
- Develop methods to address multiple colluding insiders. (Protect)
- Pursue the architecture of insider-resilient systems. (Protect)
- Pursue applications of cryptography that might limit insider threats. (Protect)
- Develop automated decoy generation (may require advances in natural language understanding). (Protect)
- Develop insider prediction techniques for users, agents, and actions. (React)

What R&D is evolutionary and what is more basic, higher risk, game changing?

Intelligent uses of authentication, existing access-control and accountability mechanisms, and behavior monitoring would generally be incremental improvements. However, in the long term, significantly new approaches are desirable.

Resources

Research, experimental testbeds, and evaluations will be essential.

Measures of success

Various metrics are needed with respect to the ability of systems to cope with insiders. Some will be generic: others will be specific to given applications and given systems. Metrics might consider the extent to which various approaches to authentication and authorization might be able to hinder insider misuse. For example, what might be the relative merits of cryptographically based authentication, biometrics, and so on, with respect to misuse, usability, and effectiveness? To what extent would various approaches to differential access controls hinder insider misuse? Detectability of insider misuse and the inviolability of audit trails would also be amenable to useful metrics.

The extent to which such localized metrics might be composable into enterprise-level metrics is a challenge of particular interest here.

To what extent can we test real systems?

- There is a strong need for realistic data for evaluation of technologies and policies that counter insider threats. This must be done operationally in a relatively noninvasive way. Testbeds are needed, as well as exportable databases of anonymized data (anonymization is generally a complicated problem).
- Red teaming is needed to identify

potential attack vectors available to insiders and to test the relevance of potential solutions.

- Some effort should be devoted to reliably simulating insider attacks and their system consequences.
- Cases of insider misuse may represent statistically rare events. Many cases of insider misuse can be expected to be unique in their motivation and execution, although there will be common modalities. Thus, special care must be devoted to understanding and accommodating the implications of rare events. Alternatively, insider misuse may be common but rarely detected or reported. If budgets are limited, choices may have to be made regarding the relative importance of improving positive and negative detection rates, and for which types of misuse cases.
- Tests involving decoys might be useful in training exercises.

References

[And2008] R. Anderson. *Security Engineering: A Guide to Building Dependable Distributed Systems*. Wiley, Indianapolis, Indiana, 2008.

[Bib1977] K.J. Biba. *Integrity Considerations for Secure Computer Systems*. Technical Report MTR 3153, The MITRE Corporation, Bedford, Massachusetts, June 1975. Also available from USAF Electronic Systems Division, Bedford, Massachusetts, as ESD-TR-76-372, April 1977.

[Bis2002] M. Bishop. *Computer Security: Art and Science*. Addison-Wesley Professional, Boston, Massachusetts, 2002.

[Bra2004] Richard D. Brackney and Robert H. Anderson. Understanding the Insider Threat: Proceedings of a March 2004 Workshop. RAND Corporation, Santa Monica, California, 2004 (http://www.rand.org/pubs/conf_proceedings/2005/RAND_CF196.pdf).

[Cap2008.1] D. Capelli, T. Conway, S. Keverline, E. Kowalski, A. Moore, B. Willke, and M. Williams. *Insider Threat Study: Illicit Cyber Activity in the Government Sector*, Carnegie Mellon University, January 2008 (http://www.cert.org/archive/pdf/insiderthreat_gov2008.pdf).

[Cap2008.2] D. Capelli, E. Kowalski, and A. Moore. Insider Threat Study: *Illicit Cyber Activity in the Information Technology and Telecommunications Sector*. Carnegie Mellon University, January 2008 (http://www.cert.org/archive/pdf/insiderthreat_it2008.pdf).

[Dag2008] Dagstuhl Workshop on Insider Threats, July 2008 (http://www.dagstuhl.de).

[FSS2008] Financial Services Sector Coordinating Council for Critical Infrastructure Protection and Homeland Security, Research and Development Committee. *Research Agenda for the Banking and Finance Sector*. September 2008 (https://www.fsscc.org/fsscc/reports/2008/RD_Agenda-FINAL.pdf). Challenge 4 of this report is Understanding the Human Insider Threat.

[HDJ2006] IT Security: Best Practices for Defending Against Insider Threats to Proprietary Data, National Defense Journal Training Conference, Arlington, Virginia. *Homeland Defense Journal*, 19 July 2006

[IAT2008] Information Assurance Technical Analysis Center (IATAC). *The Insider Threat to Information Systems: A State-of-the-Art Report*. IATAC, Herndon, Virginia, February 18, 2008.

[Kee2005] M. Keeney, D. Cappelli, E. Kowalski, A. Moore, T. Shimeali, and St. Rogers. Insider Threat Study: Computer System Sabotage in Critical Infrastructure Sectors. Carnegie Mellon University, May 2005 (http://www.cert.org/archive/pdf/insidercross051105.pdf).

[Moo2008] Andrew P. Moore, Dawn M. Cappelli, and Randall F. Trzeciak. *The "Big Picture" of IT Insider Sabotage Across U.S. Critical Infrastructures*. Technical Report CMU/SEI-2008-TR-009, Carnegie Mellon University, 2008 (http://www.cert.org/archive/pdf/08tr009.pdf). This report describes the MERIT model.

[Neu2008] Peter G. Neumann. Combatting insider misuse with relevance to integrity and accountability in elections and other applications. Dagstuhl Workshop on Insider Threats, July 2008 (http://www.csl.sri.com/neumann/dagstuhl-neumann.pdf). This position paper expands on the fuzziness of trustworthiness perimeters and the context-dependent nature of the concept of insiders.

[Noo+2008] Thomas Noonan and Edmund Archuleta. *The Insider Threat to Critical Infrastructures*. National Infrastructure Advisory Council, April 2008 (http://www.dhs.gov/xlibrary/assets/niac/niac_insider_threat_to_critical_infrastructures_study.pdf).

[Pfl2003] Charles P. Pfleeger and Shari L. Pfleeger. *Security in Computing, Third Edition*. Prentice Hall, Upper Saddle River, New Jersey, 2003.

[Ran04] M.R. Randazzo, D. Cappelli, M. Keeney, and A. Moore. Insider Threat Study: Illicit Cyber Activity in the Banking and Finance Sector, Carnegie Mellon University, August 2004 (http://www.cert.org/archive/pdf/bankfin040820.pdf).

[Sto+08] Salvatore Stolfo, Steven Bellovin, Shlomo Hershkop, Angelos Keromytis, Sara Sinclair, and Sean Smith (editors). Insider Attack and Cyber Security: Beyond the Hacker. Springer, New York, 2008.

5. Combatting Malware and Botnets

BACKGROUND

What is the problem being addressed?

Malware refers to a broad class of attack software or hardware that is loaded on machines, typically without the knowledge of the legitimate owner, that compromises the machine to the benefit of an adversary. Present classes of malware include viruses, worms, Trojan horses, spyware, and bot executables. Spyware is a class of malware used to surreptitiously track and/or transmit data to an unauthorized third party. **Bots** (short for "robots") are malware programs that are covertly installed on a targeted system, allowing an unauthorized user to remotely control the compromised computer for a variety of malicious purposes [GAO2007]. Botnets are networks of machines that have been compromised by bot malware so that they are under the control of an adversary.

Malware infects systems via many vectors, including propagation from infected machines, tricking users to open tainted files, or getting users to visit malware-propagating websites. Malware may load itself onto a USB drive inserted into an infected device and then infect every other system into which that device is subsequently inserted. Malware may propagate from devices and equipment that contain embedded systems and computational logic. An example would be infected test equipment at a factory that infects the units under test. In short, malware can be inserted at any point in the system life cycle. The World Wide Web has become a major vector for malware propagation. In particular, malware can be remotely injected into otherwise legitimate websites, where it can subsequently infect visitors to those supposedly "trusted" sites.

There are numerous examples of malware that is not specific to a particular operating system or even class of device. Malware has been found on external devices (for example, digital picture frames and hard drives) and may be deliberately coded into systems (life cycle attacks). Increasingly intelligent household appliances are vulnerable, as exemplified by news of a potential attack on a high-end espresso machine [Thu2008]. Patching of these appliances may be difficult or impossible. Table 5.1 summarizes malware propagation mechanisms.

Potentially victimized systems include end user systems, servers, network infrastructure devices such as routers and switches, and process control systems such as Supervisory Control and Data Acquisition (SCADA).

A related policy issue is that reasonable people may disagree on what is legitimate commercial activity versus malware. In addition, ostensibly legal software utilities (for example, for digital rights management [DRM]) may have unintended consequences that mimic the effects of malware [Sch2005, Hal2006].

It is likely that miscreants will develop new infection mechanisms in the future, either through discovery of new security gaps in current systems or through new exploits that arise as new communication and computation paradigms emerge.

The technical challenges are, wherever possible, to do the following:

- Avoid allowing malware onto a platform.
- Detect malware that has been installed.
- Limit the damage malware can do once it has installed itself on a platform.
- Operate securely and effectively in the presence of malware.
- Determine the level of risk based on indications of detected malware.
- Remove malware once it has been installed (remediation), and monitor and identify its source (attribution). (Remediation may sometimes be purposefully delayed on carefully monitored systems until attribution can be accomplished. Honeypots can also be useful in this regard.)

The NSA/ODNI Workshop on Computational Cyberdefense in Compromised Environments, Santa Fe, NM, August 2009, was an example of a step in this direction (http://www.c3e.info).

What are the potential threats?

Malware has significant impact in many aspects of the information age and underlies many of the topics discussed elsewhere in this document. Impacts can be single-host to networkwide, nuisance to costly to catastrophic. Negative consequences include degraded system performance and data destruction or modification. Spyware permits adversaries to log user actions (to steal user credentials and facilitate identity theft, for example), while bot malware enables an adversary to build large networks of compromised machines and amplify an adversary's digital firepower. Negative consequences of botnets and malware include spam, distributed denials of service (DDoSs), eavesdropping on traffic (sniffing), click fraud, loss of system stability, loss of confidentiality, loss of data integrity, and loss of access to network resources (for example, being identified as a bot node and then blocked by one's ISP or network administrator, effectively a DoS inflicted by one victim on another). An increasing number of websites (such as popular social networking systems, web forums, and mashups) permit user-generated content, which, if not properly checked, can allow attackers to insert rogue content that is then potentially downloaded by many users.

Beyond its nuisance impact, malware can have serious economic and national security consequences. Malware can enable adversary control of critical computing resources, which in turn may lead, for example, to information compromise, disruption and destabilization of infrastructure systems ("denial of control"), and manipulation of financial markets.

TABLE 5.1: Malware Propagation Mechanisms

Malware Propagation Mechanism	Examples
Life cycle	From the developer, either deliberate or through the use of infected development kits.
Scan and Exploit	Numerous propagating worms. May propagate without requiring action on the part of the user.
Compromised Devices	Infected USB tokens, CDs/DVDs, picture frames, etc.
Tainted File	E-mail attachment
Web	Rogue website induces user to download tainted files. (Note: Newer malware may infect victims' systems when they merely visit the rogue site, or by redirecting them to an infected site via cross-site scripting, for example)

Malware can be particularly damaging to elements of the network infrastructure. Attacks against the Domain Name System (DNS), for example, could direct traffic to rogue sites and enable a wide variety of man-in-the-middle and denial-of-service attacks. Successful attacks against DNS allow an adversary to intercept and redirect traffic, for example to rogue or spoofed servers. In addition to redirection to rogue servers, there is also the opportunity for selective or timed denial-of-service attacks; it may be easier to drop a site from DNS than to deny availability by flooding its connection. These concerns underlay the recent mandate to implement DNSSEC for the .gov domain and recommendations to implement DNSSEC for DNS root servers.

Adversaries buy and sell exploits and lease botnets in an active adversary market [Fra2007]. These botnets can be used for massive distributed attacks, spam distribution, and theft of sensitive data, such as security credentials, financial information, and company proprietary information, through sophisticated phishing attacks. The use of botnets makes attribution to the ultimate perpetrator extremely difficult. Botnets provide the adversary with vast resources of digital firepower and the potential to carry out surveillance on sensitive systems, among other threats.

Malware propagation is usually discussed in the context of enterprise and home computing. However, it also has the potential to affect control systems and other infrastructure systems. For example, the alarm systems at the Davis-Besse nuclear plant in Ohio were infected by the Slammer worm in 2003, even though these systems were supposedly immune to such an attack (the plant was not online at the time) [SF2003]. Propagating malware may have exacerbated the impact of the 2003 blackout in the northeastern United States and slowed the recovery from it. It is reasonable to assume that malware authors will target embedded systems and emerging initiatives, such as the Advanced Metering Infrastructure (AMI) for electric power.

There is also the impact associated with remediating compromised machines. From an ISP's point of view, the biggest impacts include dealing with customer support calls, purchasing and distributing antivirus (A/V) software, and minimizing customer churn. For some high-consequence government applications, an infection may even necessitate replacement of system components/hardware.

Who are the potential beneficiaries? What are their respective needs?

Malware potentially affects anyone who uses a computer or other information system. Malware remediation (cleaning infected machines, for example) is difficult in the case of professionally administered systems and beyond the technical capability of many private citizens and small office/home office (SOHO) users. Rapid, scalable, usable, and inexpensive remediation may be the most important near-term need in this topic area. Improved detection and quarantine of infected systems are also needed, as discussed below. Beneficiaries, challenges, and needs are summarized in Table 5.2.

The potential of malware to compromise confidentiality, integrity, and availability of the Internet and other critical information infrastructures is another serious concern. A real-world example would be the attacks on Estonia's cyber infrastructure via a distributed botnet in the spring of 2007 [IW2007]. That incident raised the issue of whether "cyberwar" is covered under NATO's collective self-defense mission. In the absence of robust attribution, the question remains moot. There were reports of a cyber dimension in the August 2008 conflict in the nation of Georgia, but the cyber attacks were apparently limited to denials of service against Georgian government websites and did not target cyberinfrastructure [Ant2008]. A recent malware-du-jour is Conficker, which spread initially primarily through systems that had not been upgraded with security patches, and has subsequently reappeared periodically in increasingly sophisticated versions.

The law enforcement and DoD communities are particularly interested in attribution, which, as noted above, is currently difficult.

What is the current state of the practice?

Deployed solutions by commercial antivirus and intrusion detection system/intrusion prevention system (IDS/IPS) vendors, as well as the open-source community, attempt to detect or prevent an incoming infection via a variety of vectors. A/V removal of detected malware and system reboot are currently the primary cleanup mechanisms. The fundamental challenge to this approach is that miscreants can release repacked and/or modified malware continually,

TABLE 5.2: Beneficiaries, Challenges, and Needs

Beneficiaries	Challenges	Needs
Users	Under attack from multiple malware vectors; Systems not professionally administered	User-friendly prevention, detection, containment, and remediation of malware
Administrators	Protect critical systems, maintain continuity, enterprise-scale remediation in face of explosive growth in malware variants	New detection paradigms, robust remediation, robust distribution of prevention and patches
Infrastructure Systems	Prevent accidental infection [SF 2003], address the growing challenge of targeted infection	Similar to administrator needs, but often with special constraints of legacy systems and the inability to patch and reboot at arbitrary times
ISPs	Provide continuity of service, deal with malware on more massive scale than administrators face	Defenses against propagating attacks and botnets; progress in the malware area has potential immediate impact in alleviation of these consequences
Law Enforcement	Counter growing use of malware and botnets for criminal fraud and data and identity theft	Robust attribution, advances in forensics
Government and DoD	Growing infection of defense systems, such as the Welchia intrusion into the Navy Marine Corps Intranet (NMCI) [Messmer 2003]. More recently, there have been reports of malware engineered specifically to target defense systems [LATimes08]	Share the needs of administrators, ISPs, and law enforcement

while new A/V signatures take time to produce, test, and distribute. In addition, it takes time for the user community to develop, test, and deploy patches for the underlying vulnerability that the malware is exploiting. Furthermore, the malware developers can test their software against the latest A/V versions.

Web-based A/V services have entered the market, some offering a service whereby a security professional can submit a suspicious executable to see whether it is identified as malicious by current tools. This mechanism most likely functions also as a testbed for malware developers (VirusTotal). [Vir].

Vendors of operating systems and applications have developed mechanisms for online updating and patching software for bugs, including bugs that affect security. Other defenses include antispyware, whitelists of trusted websites and machines, and reputation mechanisms.

Research in malware detection and prevention is ongoing. For example, see the Cyber-Threat Analytics project (http://www.cyber-ta.org). Also worth noting is the Anti-Phishing Working Group (APWG): http://www.antiphishing.org.

The U.S. National Institute of Standards and Technology (NIST) Security Content Automation Protocol (SCAP) is a method for using specific standards to enable automated vulnerability management, measurement, and policy compliance evaluation.

Current detection and remediation approaches are losing ground, because it is relatively easy for an adversary (whether sophisticated or not) to alter malware to evade most existing detection approaches. Given trends in malware evolution, existing approaches

(such as A/V software and system patching) are becoming less effective. For example, malware writers have evolved strategies such as polymorphism, packing, and encryption to hide their signature from existing A/V software. There is also a window of vulnerability between the discovery of a new malware variant and subsequent system patches and A/V updates. Further, malware authors also strive to disable or subvert existing A/V software once their malware has a foothold on the target system. (This is the case with a later version of Conficker, for example.) A/V software may itself be vulnerable to life cycle attacks that subvert it prior to installation. Patching is a necessary system defense that also has drawbacks. For example, the patch can be reverse engineered by the adversary to find the original vulnerability. This may allow the malware writers to refine their attacks against the unpatched systems. Much can be learned from recent experiences with successive versions of Conficker.

Specifically with respect to identity theft, which is one potential consequence of malware but may be perpetrated by other means, there is an emerging commercial market in identity theft insurance and remediation. This implies that some firms believe they have adequate metrics to quantify risk in this case.

What is the status of current research?

There is considerable activity in malware detection, capture, analysis, and defense. Major approaches include virtualization (detect/contain/capture within a virtualized environment on a particular host) [Vra2005] and honeynets (network environments, partially virtual, deployed on unused address space, that interact with malware in such a way as to capture a copy to enable further analysis) [SRI2009]. Malware is increasingly engineered to detect virtual and honeynet environments and change its behavior in response. There is industry research advancing virtual machines to the Trusted Platform Module (TPM) and hypervisor technology in hardware and software, as well as in cleanup/remediation (technically possible to do remotely in some cases, but with unclear legal and policy implications if the system owner has not given prior permission). The Department of Homeland Security has funded ongoing research in cross-domain attack correlation and botnet detection and mitigation [CAT2009]. Analysis techniques include static and dynamic analysis methods from traditional computer science.

There is considerable research into open-source IDS (SNORT and Bro) along the lines of expanding the signature base and defending these systems against adversarial intentions. Recent research has considered automatic signature generation from common byte sequences in suspicious packet payloads [Kim2004] as a countermeasure to polymorphic malware.

Significant research has been done into analysis of execution traces and similar characteristics of malware on an infected host, but we have a poor understanding of the network dimensions of the malware problem. Certain network behaviors have been observed to be important precursors to, or indicators of, malware infection. For example, DNS zone changes may predict a spam attack. Fast flux of DNS registrations (as in Conficker) may indicate that particular hosts are part of the command and control (C2) network for a large botnet. Encrypted traffic on some network ports may indicate C2 traffic to a botnet client on a given host.

Virtualization and honeynets still provide much potential in malware detection, analysis, and response, at least for the near and medium terms. For honeynets to continue to be useful, research must address issues such as:

- What features of honeynets do adversaries look for to identify them as honeynets?
- What is the ratio of "enter and retreat" to "enter and attack" in honeynets?
- How does what is actually observed in a honeynet compare with known "script kiddie" attacks and targeted malware activity in the real world?

DARPA's Self-Regenerative Systems (SRS) program developed some technology around these techniques.

Artificial diversity is transparent to correct system use but diverse from the point of view of some exploits. This has been an elusive goal, but some modest progress has been made in the commercial and research sectors. Address space randomization is now included in many operating systems; and there has been some work in the general area of system obfuscation (equivalent functionality with diverse implementation)

[Sha2004], although it has some fundamental limitations.

Emerging approaches such as behavior-based detection and semantic malware descriptions have shown promise and are deployed in commercial A/V software. However, new techniques must be developed to keep pace with the development of malware.

FUTURE DIRECTIONS

On what categories can we subdivide this topic?

For this malware and botnets topic, prevent/protect/detect/analyze/react provides a reasonable framework (see Table 5.3). **Protection** and **detection** are supported by instrumented virtualization and sandboxing environments to combat inherently secure systems, applications, and protocols. **Analysis** consists of examination of captured malware (for example, harvested on a honeynet) by IT experts in order to develop effective defenses. **Reaction** is supported by cost-effective, secure remediation that can be implemented by non-IT professionals.

What are the major research gaps?

A/V and IDS/IPS approaches are becoming less effective because malware is becoming increasingly sophisticated, and at any rate the user base (particularly consumer systems) does not keep A/V up to date. Malware polymorphism is outpacing signature generation and distribution in A/V and IDS/IPS.

Current research initiatives do not adequately address the increasing sophistication and stealth of malware, including the encryption and packing of the malicious code itself, as well as encrypted command and control channels and fast-flux DNS for botnets [Sha2008, Hol2008]. Broadly speaking, research should better understand the agility and polymorphism of malware. Automatic detection of the command and control structure of a malware sample is a significant challenge.

We do not have an adequate taxonomy of malware and botnets. It has been observed that many examples of malware are derived from earlier examples, but this avenue has not been explored as far as necessary. Progress in this area may enable, for example, defenses against general classes of malware, including as-yet unseen variants of current exemplars. A well-understood taxonomy may also support and improve attribution.

The attacker-defender relation is currently asymmetric. An attacker who develops an exploit for a particular system type will find large numbers of nearly identical exemplars of that type. Thus, it is desirable to force the adversary to handcraft exploits to individual hosts, so that the cost of developing

TABLE 5.3: Potential Approaches

Category	Definition	Potential Approaches
Prevent	Prevent the production and propagation of malware	IDS/IPS, A/V, Virtualization, Inherently secure systems
Protect	Protect systems from infection when malware is in the system's environment	IPS, A/V, Inherently secure systems
Detect	Detect malware as it propagates on networks, detect malware infections on specific systems	IDS/IPS, A/V, Virtualization, Deceptive environments
Analyze	Analyze malware's infection, propagation, and destructive mechanisms	Static and dynamic analysis, Experimentation in large-scale secured environments
React	Remediate a malware infection and identify mechanisms to prevent future outbreaks (links to the prevent category)	Updated IDS/IPS and A/V, Inherently secure systems, Thin client, Secure cloud computing paradigm

malware to compromise a large number of machines is raised significantly. Artificial diversity can address the growing asymmetry of the attacker-defender relation.

For hosts, the defenses against malware (e.g., A/V software, Windows Update, and so on) are typically part of or extensions to operating systems (OSs). This fact allows malware to easily target and disable those host-based defenses. A summary of the gaps are outlined in Table 5.4.

What are some exemplary problems for R&D on this topic?

Robust Security Against OS Exploits: Although binary-exploit malware targeting the OS is still important and worthy of incremental near-term investment, malware increasingly targets browsers and e-mail through social engineering and other mechanisms.

Protect Users from Deceptive Infections: At present, through social engineering, complexity of security controls, and rogue content injection, users can be tricked into interacting with adversary systems while thinking they are performing valid transactions, such as online banking. Research in this area should advance user education and awareness and make security controls more usable, particularly in browsers. Search engine manipulation causes the victim to go to the malware (e.g., at an infected website) rather than the malware's targeting the user (e.g., via phishing e-mail). Server-side attacks in the form of Structured Query Language

TABLE 5.4: Gaps and Research Initiatives

Identified Gap	Research Initiatives	Benefit	Time Frame
Inadequate defenses against e-mail and web malware	Human factors analysis to resist social engineering (tools, interfaces, education), Robust whitelisting	More secure present and future e-commerce	Near
Escape from virtual machines	TPM low in the hardware/ software stack	Prolongs usefulness of virtualization as a defensive strategy	Near
Difficulty of remediation	Thin client, Automatic remediation	Fast, cost-effective recovery from attack	Near
Inadequate test environments	Internet-scale emulation	Safe observation of malware spread dynamics, better containment strategies	Near
Attacker/defender asymmetry	Intentional diversity, Inherently monitorable systems	Attacker must craft attack for a large number of platforms	Medium/Long
No attack tolerance	Attack containment, Safe sandboxing, Intentional diversity	Correct operation in the presence of "subclinical" malware infection	Medium
Detection approaches losing the battle of scale	Inherently monitorable systems, Robust software whitelisting, Model-based monitoring of correct software behavior	Less space for attacker to conceal activity Detection that is generalized and scalable	Medium/Long
Inadequately understood threat	Analysis of adversary markets, Penetration of adversary communities, Containing damage of botnets while observing	Strategic view enables defensive community to take the upper hand	Long

(SQL) injection, cross-site scripting, and other methods are increasingly common ways to infect clients accessing compromised website.

Internet-scale emulation could provide game-changing breakthroughs in malware research. Being able to observe malware (specifically botnets and worms) at Internet scales without placing the real Internet in jeopardy may help identify weaknesses in the malware code and how it spreads or reacts to outside stimuli. Additionally, characteristics observed at the macro level may give us clues as to how to detect and respond to malware at the micro level. High-fidelity large-scale emulation is an important enabling capability for many of the other initiatives discussed below.

The broad area of **virtualization and honeynets** will provide much value in the near and medium terms, with respect to protection and detection approaches. Malware is becoming more adaptive, in terms of polymorphism and evasion techniques. The latter might be used to a defensive advantage. If malware is designed to be dormant if it detects that it is in a virtual machine or in a honeynet environment, active deception on the part of the defender (making production systems look like virtual systems and production networks look like honeynets, and vice versa; changing virtual and real systems very rapidly; or even the use of an analog to a "screen saver" that toggles a computer from real to honeynet when the user is not actively using it) may prove useful. The general research question is how "deception" can be best leveraged by defenders.

There are concerns about the limitations of these approaches. Even a correctly functioning hypervisor is inadequate in case of some flaws in the guest OS, for example. Also, highly sophisticated malware is likely to be able to escape current-generation virtual environments. Improved hardware architecture access mechanisms will maintain the effectiveness of these approaches to some degree. However, additional research is needed on techniques that seize the strategic low ground within our computing systems and also separate the security functions from other functionality. The key insight is that our detection methods and instrumentation must reside lower in the hardware/software stack than the malware. Otherwise, the malware controls the defenders' situational awareness, and the defenders have no chance. Recent research injecting vulnerabilities into hardware designs suggests disturbing possibilities for the future on this front.

Collaborative detection may involve privacy-preserving security information sharing across independent domains that may not have an established trust relationship. We may share malware samples, metadata of a sample, and experiences. A repository of active malware may accelerate research advances but raises security concerns in its own right, and access must be carefully controlled according to a policy that is difficult to define. Moreover, sharing malware may be illegal, depending on the business of the entity.

Collaborative detection supports an identified need in the situational understanding topic area. In particular, the detection, quarantine, and remediation of botnet assets is a major overlap between the research needs for malware and those of situational understanding (Section 8). Network-level defenses must come online to supplement host-level defenses. For example, we require better identification of bad traffic at the carrier level. This presents challenges in scale and speed.

Thin-client technology has been proposed in the past. In this model, the user's machine is stateless, and all files and applications are distributed on some network (the terminology "in the cloud" is occasionally used, although there are also parallels with traditional mainframe computing). If we can make the distributed resources secure, and that is itself a big question, the attacker options against user asset are greatly reduced, and remediation is merely a question of restarting. The long-term research challenges toward this secure cloud computing paradigm are securing the distributed resource base and making this base available to the authenticated and authorized user from any location, supported by a dedicated, integrated infrastructure.

Remediation of infected systems is extremely difficult, and it is arguably

impossible to assert that a previously infected system has in fact been thoroughly cleansed. In particular, systems may be infected with rootkits, which come in many forms, from user level to kernel level rootkits. More recently, hardware virtual machine (HVM) rootkits have been proposed, which load themselves into an existing operating system, transforming it into a guest OS controlled by the rootkit [Dai2006]. We require advances in remediation, built-in diagnostic instrumentation, and VM introspection that provides embedded digital forensics to deal with these threats.

Containment technology (which includes TPM approaches mentioned previously) is promising but needs further work. An interesting goal is to tolerate malware (for example, safely doing a trusted transaction from a potentially untrusted system). Another goal is to have a "safe sandbox" for critical transactions (in contrast to current sandboxing environments that typically seek to contain the malware in the sandbox). A final issue is whether large systems can achieve their goal while tolerating a residual level of ongoing compromise within their components and subsystems. Generally, the research agenda should recognize that malware is part of the environment, and secure operation in the presence of malware is essential.

Development of **inherently secure, monitorable, and auditable systems** has presented a significant challenge. In general, this is a medium- to long-term research area. Short-term work in trusted paths to all devices may reduce the risk of, for example, key logging software. In the short term, we require advances in authenticated updates, eventually evolving systems that are immune to malware. Advances in this area relate to the scalable trustworthy systems topic in Section 1.

A longer-term research challenge is to develop systems, applications, and protocols that are inherently more secure against malware infection and also easier to monitor in a verifiable way (in effect, to reduce the space in which malware can hide within systems). In particular, hardware-based instrumentation that provides unbiased introspection for and unimpeded control of COTS computing devices, while being unobservable by the malware, may help enable embedded forensics and intrinsically auditable systems.

Artificial diversity can take many forms: the code is different at each site, the location of code is different, system calls are randomized, or other data is changed. It may be worth researching (both in terms of practicality and economics) how to randomize instruction sets, operating systems, and libraries that are loaded from different system reboots. A difficult end goal would be to develop systems that function equivalently for correct usage but are unique from an attack standpoint, so an adversary must craft attacks for individual machines. Artificial diversity is just one approach to **changing the attacker-defender asymmetry**, and novel ideas are required.

Not enough is being done in **threat analysis**. In any case, the nature of the threat changes over time. One interesting avenue of research is economic analysis of adversary markets. Attackers sell malware exploits (and also networks of infected machines, or botnets). The price fluctuations may permit analysis of adversary trends and may also enable definition of metrics as to the effectiveness of defenses. Related to the economic approach is research into making malware economically less attractive to adversaries (for example, by much better damage containment, increasing the effectiveness of attribution, limiting the number of systems that can be targeted with a given exploit, and changing existing laws/policies so that the punishments reflect the true societal cost of cybercrime).

What R&D is evolutionary, and what is more basic, higher risk, game changing?

In the near term, we are in a defensive struggle, and R&D should continue in the promising areas of virtualization and honeynets. We require near-term advances in remediation to address the serious and increasing difficulty of malware cleanup, particularly on end-user systems. Research in the area of attack attribution in the near and medium terms can aid the policing that is necessary on the Internet. Mechanisms to share data from various kinds of malware attacks are currently lacking, as well. The problems faced by researchers

in this domain range from privacy concerns, legal aspects of data sharing, and the sheer volume of data itself. Research in generating adequate metadata and provenance is required to overcome these hurdles.

Techniques to capture and analyze malware and propagate defenses faster are essential in order to contain epidemics. Longer-term research should focus on inherently secure, monitorable, and auditable systems. Threat analysis and economic analysis of adversary markets should be undertaken in pilot form in the near term, and pursued more vigorously if they are shown to be useful.

Measures of success

We require baseline measurements of the fraction of infected machines at any time; success would be a reduction in this fraction over time.

Some researchers currently track the emergence of malware. In this way, they are able to identify trends (for example, the number of new malware samples per month). A reversal of the upward trend in malware emergence would indicate success.

Time between malware capture and propagation of defense (or, perhaps more appropriately, implementation of the defense on formerly vulnerable systems) tracks progress in human and automated response time.

With reference to the repository, we may define a minimal set of exemplars that must be detected in order to claim effectiveness at some level.

We can define measures of success at a high level by answering the following questions and tracking the answers over time:

- How many machines do we know about that serve malware?
- What is the rate of emergence of new malware?
- Since spam is a primary botnet output, what fraction of e-mail is spam?
- What is the industry estimate of hosts serving malware?
- What is the trend in malware severity (on a notional continuum, say from nuisance to adware, spyware, bot capture)?
- What fraction of known attacks is successful, and what fraction is thwarted?

We may also consider cost-based measures (from the defender point of view), such as:

- What is our cost of searching for malware propagators?
- What is the cost to identify botnets and their bot command and control infrastructures?
- What is the cost to increase sharing of malware host lists?

Economic analysis of adversary markets may allow definition of metrics as to effectiveness of particular defenses.

It would be beneficial to have reliable metrics that estimate the vulnerability of particular systems to corruption by malware, and how well they are able to withstand other kinds of malware-enabled attacks, such as DDoS attacks. Similarly, metrics that suggest the benefits that will accrue with the use of particular malware prevention or remediation strategies would be helpful.

What needs to be in place for test and evaluation?

Beyond reverse engineering of malware, the most effective studies of malicious code have taken place on network testbeds. These testbeds have included simple virtual machines "networked" on an analyst's computer, testbeds consisting of tens or hundreds of real (nonvirtualized) nodes, such as DETER [DET], and simulated networks created within network simulation tools. The research community has yet to approach studies of malware in Internet-scale emulated environments. The infrastructure and tools do not currently exist to build emulation environments on the order of 10,000,000 nodes or more.

As malware sophistication improves to include detection of virtual environments, the realism of the virtualization environment (for example, virtual machine or honeynet) testbed presents a challenge.

Tools and environments to study malware need to evolve as the malware evolves. In particular, the community

currently does not have testbeds for hardware/firmware-based malware.

The tools and infrastructure required to adequately harden a test environment are research problems in their own right. Testbeds to study malware are specific to this application. The testbed should not be discernible as a test environment, even to sophisticated malware.

The community requires an up-to-date, reliably curated malware repository for research purposes. Limited repositories exist at present, but they are not available to the research community. Another desirable resource would be a shared honeynet, which would allow learning malware behavior. Current honeynets are run mostly on an ad hoc basis by individual groups. Legal and regulatory issues inhibit meaningful sharing, however.

Internet-scale emulation would permit realistic testing of defenses and their dynamic interaction with malware outbreaks. Observation at this level would provide a view of worm and botnet spread and operation never seen before.

To what extent can we test real systems?

It is possible to test defenses for efficacy on real systems. Experiments can be conceived in which real and emulation networks are exposed to public networks, with and without particular defenses. However, rapid automated configuration and propagation of defenses must first be thoroughly demonstrated on emulated systems.

References

[Ant2008] A.M. Antonopoulos. Georgia cyberwar overblown. *Network World*, August 19, 2008 (http://www.pcworld.com/businesscenter/article/150021/georgia_cyberwar_overblown.html).

[CAT2009] Conference for Homeland Security 2009 (CATCH '09), Cybersecurity Applications and Technology, March 3–4, 2009. The IEEE proceedings of this conference include relevant papers on detection and mitigation of botnets, as well as correlation and collaboration in cross-domain attacks, from the University of Michigan and Georgia Tech, as well as Endeavor, HBGary, Milcord, and Sonalyst (among others).

[Dai2006] Dino Dai Zovi, Vitriol: Hardware virtualization rootkits. In *Proceedings of the Black Hat USA Conference*, 2006.

[DET] Cyber-DEfense Technology Experimental Research laboratory Testbed (DETERlab) (http://www.isi.edu/deter/).

[Fra2007] J. Franklin, V. Paxson, A. Perrig, and S. Savage. An inquiry into the nature and causes of the wealth of Internet miscreants. *Proceedings of ACM Computer and Communications Security Conference*, pp. 375-388, October 2007.

[GAO2007] *CYBERCRIME: Public and Private Entities Face Challenges in Addressing Cyber Threats*. Report GAO-07705, U.S. Government Accountability Office, Washington, D.C., July 2007.

[Hal2006] J.A. Halderman and E.W. Felten. Lessons from the Sony CD DRM episode. In *Proceedings of the 15th USENIX Security Symposium*, August 2006.

[Hol2008] T. Holz, C. Gorecki, K. Rieck, and F. Freiling. In *Proceedings of the 15th Annual Network & Distributed System Security (NDSS) Symposium,* February 2008.

[Kim2004] Hyang-Ah Kim and Brad Karp, Autograph: Toward automated, distributed worm signature detection, In *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[IW2007] L. Greenemeier. Estonian attacks raise concern over cyber 'nuclear winter.' *Information Week*, May 24, 2007 (http://www.informationweek.com/news/internet/showArticle.jhtml?articleID=199701774).

[LAT2008] J.E. Barnes. Cyber-attack on Defense Department computers raises concerns. *Los Angeles Times*, November 28, 2008 (http://www.latimes.com/news/nationworld/iraq/complete/la-na-cyberattack28-2008nov28,0,230046.story).

[Mes2003] Ellen Messmer. Welchia Worm Nails Navy Marine Corps, Network World Fusion, August 19, 2003. (http://pcworld.com/article/112090/welchia_worm_nails_navy_marine_corps.html).

[Pou2003] Kevin Poulsen. Slammer worm crashed Ohio nuke plant network. *SecurityFocus*, August 19, 2003 (http://www.securityfocus.com/news/6767).

[Sha2004] H. Shacham, M. Page, B. Pfaff, E.-J. Goh, N. Modadugu, and D. Boneh. On the effectiveness of address-space randomization. In *Proceedings of the 11th ACM Computer and Communications Security Conference*, Washington, D.C., pp. 298-307, 2004.

[Sha2008] M. Sharif, V. Yegneswaran, H. Saidi, P. Porras, and W. Lee. Eureka: A framework for enabling static malware analysis. In *Proceedings of the 13th European Symposium on Research in Computer Security (ESORICS)*, Malaga, Spain, pp. 481-500, October 2008.

[Sch2005] Bruce Schneier. Real story of the rogue rootkit. *Wired*, November 17, 2005 (http://www.wired.com/politics/security/commentary/securitymatters/2005/11/69601).

[SRI2009] SRI Cyber-Threat Analytics (http://www.cyber-ta.org/) and Malware Threat Center (http://mtc.sri.com). For example, see analyses of Conficker.

[Thu2008] R. Thurston. Coffee drinkers in peril after espresso overspill attack. *SC Magazine*, June 20, 2008 (http://www.scmagazineuk.com/coffee-drinkers-in-peril-after-espresso-overspill-attack/article/111458).

[Vir] Virus Total (http://www.virus-total.com).

[Vra+2005] M. Vrable, J. Ma, J. Chen, D. Moore, E. Vandekieft, A. Snoeren, G. Voelker, and S. Savage. Scalability, fidelity and containment in the Potemkin virtual honeyfarm. *ACM SIGOPS Operating Systems Review*, 39(5):148-162, December 2005 (SOSP '05).

6. Global-Scale Identity Management

BACKGROUND

What is the problem being addressed?

Global-scale identity management concerns identifying and authenticating entities such as people, hardware devices, distributed sensors and actuators, and software applications when accessing critical information technology (IT) systems from anywhere. The term global-scale is intended to emphasize the pervasive nature of identities and implies the existence of identities in federated systems that may be beyond the control of any single organization. This does not imply universal access or a single identity for all purposes, which would be inherently dangerous. In this context, global-scale identity management encompasses the establishment of identities, management of credentials, oversight and accountability, scalable revocation, establishment and enforcement of relevant policies, and resolution of potential conflicts. To whatever extent it can be automated, it must be administratively manageable and psychologically acceptable to users. It must, of course, also be embedded in trustworthy systems and be integrally related to authentication mechanisms and authorization systems, such as access controls. It also necessarily involves the trustworthy binding of identities and credentials. It is much broader than just identifying known individuals. It must scale to enormous numbers of users, computer systems, hardware platforms and components, computer programs and processes, and other entities.

Global-scale identity management is aimed specifically at government and commercial organizations with diverse interorganizational relationships that today are hampered by the lack of trustworthy credentials for accessing shared resources. In such environments, credentials tend to proliferate in unmanageable ways. Identity management within single organizations can benefit from—and needs to be compatible with—the global-scale problem.

Our concern here is mainly the IT-oriented aspects of the broad problems of identity and credential management, including authentication, authorization, and accountability. However, we recognize that there will be many trade-offs and privacy implications that will affect identity management. In particular, global-scale identity management may require not only advances in technology, but also open standards, social norms, legal frameworks, and policies for the creation, use, maintenance, and audit of identities and privilege information (e.g., rights or authorizations). Clearly, managing and coordinating people and other entities on a global scale also raises many issues relating to international laws and regulations that must be considered. In addition, the question of when identifying information must be provided is fundamentally a policy question that can and should be considered. In all likelihood, any acceptable concept of global identity management will need to incorporate policies governing release of identifying information. Overall, countless critical systems and services require authenticated authorization for access and use,

and global-scale identity management will be a critical enabler of future IT capabilities. Furthermore, it is essential to be able to authorize on the basis of attributes other than merely supposed identities. Identity management needs to be fully integrated with all the systems into which it is embedded.

Identity management systems must enable a suite of capabilities. These include control and management of credentials used to authenticate one entity to another, and authorization of an entity to adopt a specific role and assert properties, characteristics, or attributes of entities performing in a role. Global-scale identity management must also support nonrepudiation mechanisms and policies; dynamic management of identities, roles, and properties; and revocation of properties, roles, and identity credentials. Identity management systems must provide mechanisms for two-way assertions and authentication handshakes building mutual trust among mutually suspicious parties. All the identities and associated assertions and credentials must be machine and human understandable, so that all parties are aware of the identity interactions and relationships between them (e.g., what these credentials are, who issued them, who has used them, and who has seen them). The lifetimes of credentials may exceed human lifetimes in some cases, which implies that prevention of and recovery from losses are particularly difficult problems.

What are the potential threats?

Identification and authentication (I&A) systems are being attacked on many fronts by a wide range of potential attackers with diverse motivations, within large-scale organizations and across multiple organizations. Insider and outsider misuses are commonplace. Because of the lack of adequate identity management, it is often extremely difficult to identify the misusers. For example, phishing attacks have become a pervasive problem for which identifying the sources and the legitimacy of the phishers and rendering them ineffective where possible are obvious needs.

Identity-related threats exist throughout the development cycle and the global supply chain, but the runtime threats are generally predominant. Misuse of identities by people and misuse of flawed authentication by remote sites and compromised computers (e.g., zombies) are common. The Internet itself is a source of numerous collateral threats, including coordinated, widespread denial-of-service attacks, such as repeated failed logins that result in disabling access by legitimate users. Various threats arise when single-sign-on authentication of identities occurs across boundaries of comparable trustworthiness. This is likely to be a significant concern in highly distributed, widespread system environments. Additional threats arise with respect to the misuse of identities and authentication, especially in the presence of systems that are not adequately trustworthy. Even where systems have the potential for distinguishing among different roles associated with different individuals and where fine-grained access controls can be used, operational considerations and inadequate user awareness can tend to subvert the intended controls. In particular, threats are frequently aimed at violations of integrity, confidentiality, and system survivability, as well as denial-of-service attacks.

Threats described in other topic areas can also affect global-scale identity management, most notably defects in trustworthy scalable systems. In addition, defects in global-scale identity management can have negative impacts on provenance and attack attribution.

Who are the potential beneficiaries? What are their respective needs?

Governmental agencies, corporations, institutions, individuals, and particularly the financial communities [FSSCC 2008] would benefit enormously from the existence of pervasive approaches to global identity management, with greater convenience, reduction of administrative costs, and possibilities for better oversight. Users could benefit from the decreased likelihood of impersonation, identity and credential fraud, and untraceable misuse. Although the needs of different individuals and different organizations might differ somewhat, significant research in this area would have widespread benefits for all of them.

What is the current state of the practice?

There are many current approaches to identity management. Many of these are not yet fully interoperable with other required services, not scalable, only single-use, or limited in other ways. They do, however, collectively exhibit pointwise examples that can lead toward enabling a global-scale identity

management framework. Examples of existing approaches include the following:

- Personal ID and authentication. Shibboleth is a standards-based, open-source software system for single sign-on across multiple websites. (See http://shibboleth.internet2.edu.) Also of interest are Card Space, Liberty Alliance, SAML, and InCommon (all of which are federated approaches, in active use, undergoing further development, and evolving in the face of various problems with security, privacy, and usability).
- The Homeland Security Presidential Directive 12 (HSPD-12) calls for a common identification standard for federal employees and contractors. An example of a solution in compliance with HSPD-12 is the DoD Common Access Card (CAC).

Various other approaches such as the following could play a role but are not by themselves global-scale identity solutions. Nevertheless, they might be usefully considered. Open ID provides transitive authentication, but only minimal identification; however, trust is inherently not transitive, and malicious misuse is not addressed. Medical ID is intended to be HIPAA compliant. Enterprise Physical Access is representative of token-based or identity-based physical access control systems. Stateless identity and authentication approaches include LPWA, the Lucent Personalized Web Assistant. OTP/VeriSign is a symmetric key scheme. Biometrics could potentially be useful as part of the authentication process, but most biometric technologies currently have various potential implementation vulnerabilities, such as fingerprint readers being fooled by fake gelatin fingers. Credit cards, debit cards, smart cards, user-card-system authentication, and chip and PIN have all experienced some vulnerabilities and various misuses. Per-message techniques such as DKIM (DomainKeys Identified Mail), authenticating e-mail messages, PGP, and S/MIME are also worth considering—especially for their limitations and development histories.

It is desirable to learn from the relative shortcomings of all these approaches and any experience that might be gained from their deployment. However, for the most part, these sundry existing identity management concepts do not connect well with each other. Forming appropriate and effective, semantically meaningful connections between disparate identity management systems presents a significant challenge. Given a future with many competing and cooperating identity management systems, we must develop a system of assurance for the exchange of identity credentials across identity management systems, and principled means to combine information from multiple identity management systems as input to policy-driven authorization decisions. The threats noted above are poorly addressed today.

What is the status of current research?

Currently, there are several major initiatives involving large-scale identity management, including a government-wide E-Authentication initiative, the Defense Department's Common Access Card, and public key infrastructure for the Global Information Grid. These are not research directions, but exhibit many problems that can motivate future research. However, none of these can scale to the levels required without substantial problems regarding federation of certification authorities and delays in handling revoked privileges. Moreover, although it is perhaps a minor consideration today, the existing standard and implementations are based on public-key cryptography that could eventually be susceptible to attack by quantum computers.

Considerable research exists in policy languages, trust negotiation, and certificate infrastructures that have not yet been tried in practice. Research strategies to achieve a strong I&A architecture for the future include large-scale symmetric key infrastructures with key distribution a priori, federated systems of brokers to enable such a system to scale, strategies for scaling symmetric creation of one-time pads, schemes of cryptography not reliant on a random oracle, and other schemes of cryptography not susceptible to attack by quantum computers (which seems possible, for example, with lattice-based cryptography). The series of IDtrust symposia at NIST summarize much work over the past 9 years [IDT2009], including three papers from the 2009 symposium from an ongoing collaborative I3P project on identity management. On the other hand, relatively little work has been done on avoiding monolithic trusted roots, apart from systems such as Trusted Xenix. There is also not

enough effort devoted to trustworthy bindings between credentials and users. Biometrics and radio frequency identification (RFID) tags both require such binding. However, by no means should research on potential future approaches be limited to these initial ideas.

FUTURE DIRECTIONS

On what categories can we subdivide the topic?

Two categories seem appropriate for this topic area, although some of the suggested research areas may require aspects of both categories:

- Mechanisms (e.g., for authentication, attribution, accountability, revocation, federation, usable user interfaces, user-accessible conceptual models, presentation, and evaluations thereof).
- Policy-related research (e.g., privacy, administration, revocation policies, international implications, economic, social and cultural mores, and policies relating to the effective use of the above mechanisms)

As is the case for the other topics, the term "research" is used here to encompass the full spectrum of R&D, test, evaluation, and technology transfer. Legal, law enforcement, political, international, and cultural issues are cross-cutting for both of these bins and need to be addressed throughout.

Mechanisms for enhancing global identity management (with some policy implications) include the following:

- Federated bilateral user identity and credential management on a very large scale, to facilitate interoperability among existing systems.
- Efficient support for management of identities of objects, processes, and transactions on a very large scale.
- Flexible management of identities (including granularity, aliases, proxies, groups, and associated attributes).
- Support for multiple privacy and cross-organization information exposure requirements, lightweight aliasing, and unlinking.
- Effective presentation of specific attributes: multiple roles, multiple properties, effective access rights, transparency of what has and has not been revealed.
- Enabling rapidly evolving and newly created attributes, such as value associated with identifiers.
- Timely revocation of credentials (altering or withdrawing attributes).
- Avoidance of having to carry too many certificates versus the risks of single-sign-on authentication that must be trustworthy despite traversing untrustworthy systems.
- Long-term implications of cryptographically based approaches, with respect to integrity, spoofability, revocation when compromised, accountability, credential renewals, problems that result from system updates, and so on.
- Identity management for nonhuman entities such as domain names, routers, routes, autonomous systems, networks, and sensors.

Note that merely making SSL client certificates work effectively in a usable way might be a useful initial step forward.

Policies for enhancing global identity management (some of which have mechanism implications) include the following.

- Risk management across a spectrum of risks. This is tightly coupled with authorization. Game-theoretical analyses might be useful.
- Trust or confidence in the interactions (untrustworthy third parties; what happens when your credentials get stolen or the third party disappears).
- User acceptance: usability, interoperability, costs; fine-grained attribute release and presentation to users.
- Explicating the structure, meaning, and use of attributes: semantics of identity and attribute assertions.
- Commercial success and acceptance: usability, interoperability, costs, sustainable economic models; presentation to users.
- Accommodating international implications that require special

consideration, such as seemingly fundamental differences in privacy policies among different EU nations, the United States, and the rest of the world.

- Compensating for possible implications of new approaches that enable new types of transactions and secondary uses that were not initially anticipated.
- Understanding the implications of quantum computing and quantum cryptography, and exploring the possibilities of global identity management without public-key cryptography or with quantum-resistant public-key cryptography.

Table 6.1 provides an oversimplified summary of the two categories.

What are the major research gaps?

A key gap in identity management is the lack of transparent, fine-grained, strongly typed control of identities, roles, attributes, and credentials. Entities must be able to know and control what identity-related information has been provided on their behalf. Entities must be able to present credentials for identities, roles, and attributes—independently but consistently interrelated, relative to specific needs. For example, why should a liquor store clerk be able to view a person's address and other personal details on a driver's license when determining whether that person is at least 21, or, worse yet, to swipe a card with unknown consequences? Services should be able to validate role or property credentials for some situations without requiring explicit identity as well. Entities and services must also be able to select appropriate levels of confidence and assurance to fit their situation. In addition, secondary reuse of credentials by authorizing entities must be effectively prevented. Some sort of mutual authentication should be possible whenever desirable. That is, a bidirectional trusted path between the authenticatee and the authenticator may be needed in some cases.

Major gaps include the following:

- Existing systems tend to authenticate only would-be identities of users, not transactions, applications, systems, communication paths, hardware, individual packets, messages, and so on.
- Containment, detection, and remediation are poorly addressed, particularly following misuse of identities, authentication, and authorization.
- Maintaining consistency of reputations over time across identities is extremely difficult. However, carefully controlled mechanisms to revoke or otherwise express doubts about such reputations are also needed.
- Past efforts to impose national standards for identity management have met considerable resistance (as in Australia and the United Kingdom).
- There is a serious lack of economic models that would underscore the importance of global-scale identity management and lead to coherent approaches.
- There is also a serious lack of understanding of cultural and social implications of identity, management authentication, and privacy among most citizens.

TABLE 6.1: Some Illustrative Approaches

Category	Definition	Potential Approaches
Mechanisms	Identity- and attribute-based systems implementing authentication, authorization, accountability	Globally trustworthy identities, cryptographic and biometric authentication, secure bindings to entities, distributed integrity
Policies	Rules and procedures for enforcing identity-based controls, using relevant mechanisms	Broadly based adversary detection systems that integrate misuse detection, network monitoring, distributed management

Achieving the goal of open, globally accepted standards for identifying individuals, system components, and processes is difficult and will take considerable coordination and cooperation between industry and governments. Global-scale identity management is a hard problem for a number of reasons, including standardization, scale, churn, time criticality, mitigation of insider threats, and the prospect of threats such as quantum computing to existing cryptographic underpinnings. Maintaining the anonymity of personal information unless explicitly required is another challenge. In addition, determining how system processes or threads should be identified and privileged is an even more complex and daunting undertaking. Part of the challenge is to distinguish between the user and the subjects executing on his or her behalf. Finally, although sensor networks and radio frequency identification (RFID) have tremendous utility, their current vulnerabilities and the desired scale of future deployment underscore the need to address the hard challenges of identity management on a global scale.

Resources

Short-term gains can be made, particularly in prototypes and in the policy research items noted in the Background section above. In particular, the intelligent use of existing techniques and implementations would help. However, serious effort needs to be devoted to long-term approaches that address inherent scalability, trustworthiness, and resistance to cryptanalytic and systemic attacks, particularly in federated systems in which trustworthiness can not be assured.

Measures of success

Ideally, any system for identification, authentication, and access control should be able to support hundreds of millions of users with identity-based or role-based authentication. IDs, authentication, and authorization of privileges may sometimes be considered separately, but in any case must be considered compatibly within a common context. An identifier declares who a person is and may have various levels of granularity and specificity. Who that person is (along with the applicable roles and other attributes, such as physical location) will determine the privileges to be granted with respect to any particular system policy. The system should be able to handle millions of privileges and a heavy churn rate of changes in users, devices, roles, and privileges. In addition, each user may have dozens of distinct credentials across multiple organizations, with each credential having its own set of privileges. It should be possible to measure or estimate the extent to which incremental deployment of new mechanisms and new policies could be implemented and enforced. Revocation of privileges should be effective for near-real-time use. Measurable metrics need to encompass all these aspects of global identity management. Overall, it should be extremely difficult for any national-level adversary to spoof a critical infrastructure system into believing that anyone attempting access is anything other than the actual adversary or adversaries.

Some of the possibly relevant metrics might involve the following considerations:

- Interoperability. How many systems might be integrated? What efficiency can result as scopes of scalability increase?
- Bilateral identity management. How many identities might be handled? What are the risks?
- Efficiency of identity transactions at global scale. For example, what is the end-to-end minimum time to process various types of transactions?
- Revocation. What are the time delays for expected propagation as the global scale increases?
- Value metrics. What are the short-term and long-term values that might result from various approaches?
- Privacy metrics. For example, how easily can behavior analysis or pseudonymous profiling be used to link multiple identities?
- Risk management metrics. What are the risks associated with the above items?

What needs to be in place for test and evaluation?

Federated solutions will require realistic testbeds for test and evaluation of global identity management approaches. Universities would provide natural environments for initial experimentation and might, under controlled circumstances, enable larger-scale collaborations.

Numerous opportunities will exist for formal analysis of algorithms and prototypes, especially as they scale up to federated solutions. These should complement any testing.

To what extent can we test real systems?

Today's test and evaluation are rather ad hoc and leave beta testing to user communities. Test criteria, scalability, robustness, and cost need to be considered. Some things can be tested; others require different kinds of analysis, including large-scale simulations and formal methods. Scalability is needed with respect to the number of organizational and multi-organizational requirements, and the number of organizations, not just the number of people. Testing is only part of what is necessary. Federated algorithms need some formal analyses with respect to their consistency, security, and reliability. Experiences with failed or ineffective attempts in the past must be reflected in new directions. As is often the case, sharing of such experiences is difficult. So are multi-institutional testbeds and experiments. Incentives are needed to facilitate sharing of experiences relating to vulnerabilities and exploits. Algorithmic transparency is needed, rather than closely held proprietary solutions.

Approaches to test markets require specific attention to usefulness and usability and to cost-effectiveness. Possible test markets include virtual environments such as World of Warcraft or Second Life and real-world environments such as banking, financial services, eBay, the Department of Energy, Department of Veterans Affairs, federated hospitals, and Las Vegas casinos. Realistic testbeds require realistic incentives such as minimizing losses, ability to cope with large-scale uses, ease of evaluation, and trustworthiness of the resulting systems—including resilience to denials of service and other attacks, overall system survivability, and so on.

References

[FSS2008] Financial Services Sector Coordinating Council for Critical Infrastructure. Protection and Homeland Security, Research and Development Committee. Research Agenda for the Banking and Finance Sector. September 2008, (https://www.fsscc.org/fsscc/reports/2008/RD_Agenda-FINAL.pdf).

[IDT2009] 8th Symposium on Identity and Trust on the Internet (IDtrust 2009), NIST, April 14-16, 2009 (http://middleware.internet2.edu/idtrust). The website contains proceedings of previous years' conferences. The 2009 proceedings include three papers representing team members from the I3P Identity Management project (which includes MITRE, Cornell, Georgia Tech, Purdue, SRI, and the University of Illinois at Urbana-Champaign).

7. Survivability of Time-Critical Systems

BACKGROUND

What is the problem being addressed?

Survivability is the capability of a system to fulfill its mission, in a timely manner, in the presence of attacks, failures, or accidents [Avi+1994, Ell+1999, Neu2000]. It is one of the attributes that must be considered under trustworthiness, and is meaningful in practice only with respect to well-defined mission requirements against which the trustworthiness of survivability can be evaluated and measured.

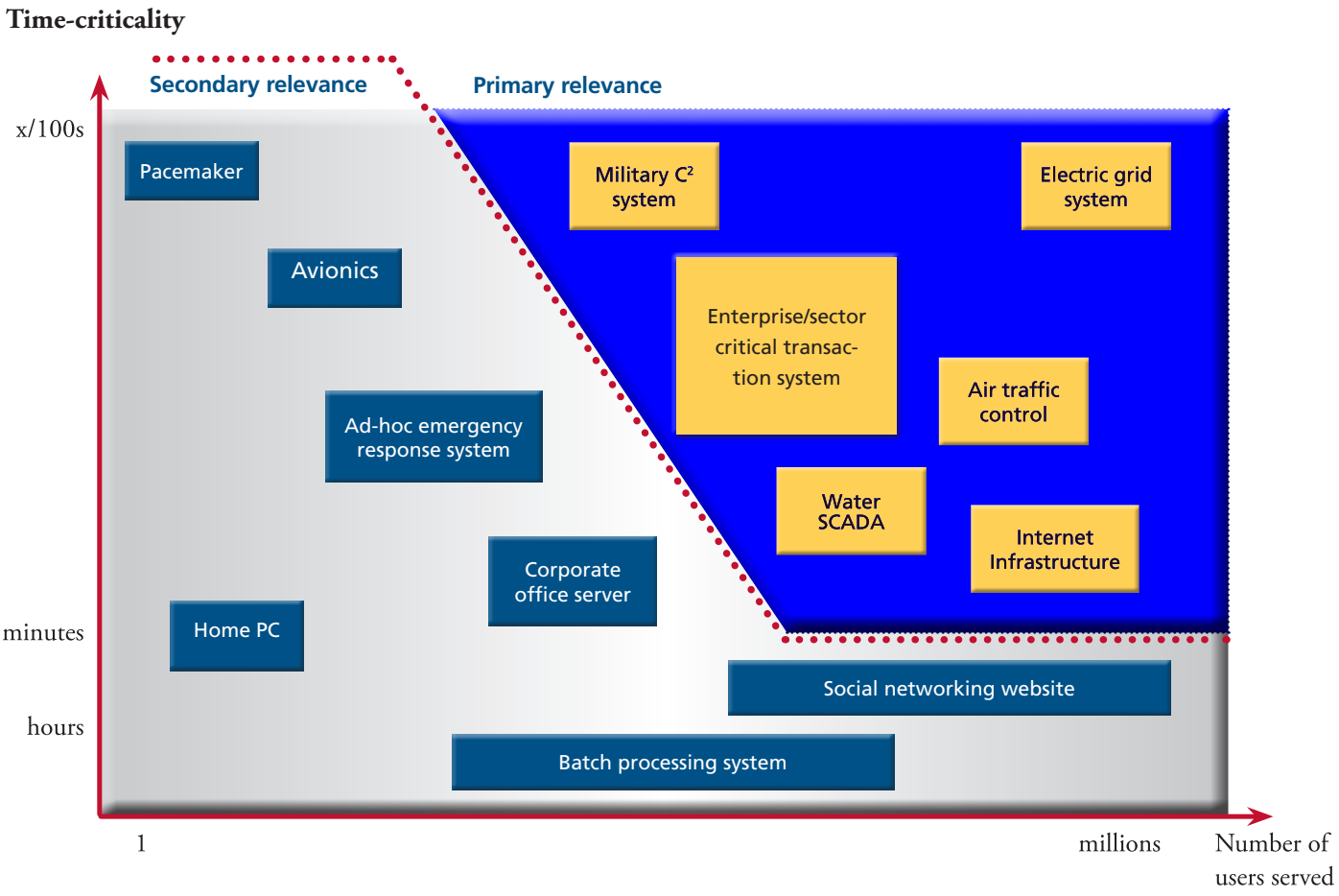
Time-critical systems, generally speaking, are systems that require response on non-human timescales to maintain survivability (i.e., continue to operate acceptably) under relevant adversities. In these systems, human response is generally infeasible because a combination of the complexity of the required analysis, the unavailability and infeasibility of system administrators in real time, and the associated time constraints. This section uses the following definition:

With respect to survivability, a **time-critical system** is a system for which faster-than-human reaction is required to avoid adverse mission consequences and/or system instability in the presence of attacks, failures, or accidents.

Of particular interest here are systems for which impaired survivability would have large-scale consequences, particularly in terms of the number of people affected. Examples of such systems include electric power grids and other critical infrastructure systems, regional transportation systems, large enterprise transaction systems, and Internet infrastructure such as routing or DNS. Although impaired survivability for some other types of systems may have severe consequences for small numbers of users, they are not of primary relevance to this topic. Examples of such systems are medical devices, individual transportation systems, home desktop computers, and isolated embedded systems. Such systems are not always designed for an adequate level of survivability, but the problem is less challenging to address for them than for large and distributed systems. However, common-mode failures of large numbers of small systems (for example, a vulnerability in a common type of medical device) could have large-scale consequences. (Note that personal systems are not actually ignored here, in that certain major advances in survivability of large-scale time-critical systems may be applicable to smaller systems.)

Time criticality is a central property to be considered. It connects directly to the "faster-than-human" aspect of the above definition of survivability. In some systems, failure to fulfill a mission for even fractions of a second could have severe consequences. In other types of systems, downtime for several minutes could be acceptable. In some other systems, system stability could be threatened if upsets are not handled on faster-than-human timescales. See Figure 7.1 for examples

Figure 7.1: Examples of Systems With Different Time-Criticality Requirements and Different User Populations



of systems categorized with respect to relative time criticality and size of the user population they serve. The systems on the right side of the diagonal line are considered in primary scope for this discussion, while systems to the left of the line are of secondary interest, as indirect beneficiaries.

What are the potential threats?

As noted in the definition of survivability, the threats include system attacks, failures, and accidents. Rather than enumerate a long list, we refer throughout to "all relevant adversities" for which survivability is required.

Who are the potential beneficiaries? What are their respective needs?

Beneficiaries include the ultimate end users of critical infrastructure systems (the public), system owners and operators, system developers and vendors, regulators and other government bodies, educators and students, standards bodies, and so on. These categories of beneficiaries have very different needs. End users need to have a working system whenever they need to use it (availability), and they need the system to continue working correctly once they have started using it (reliability). System owners have many additional needs; for example, they need to have situational awareness so that they can be warned about potential problems in the system and manage system load, and they need to be able to react to an

incident and to recover the system and restore operations.

What is the current state of practice?

At present, IT systems attempt to maximize survivability through replication of components, redundancy of information (e.g., error-correcting coding), smart load sharing, journaling and transaction replay, automated recovery to a stable state, deferred committing for configuration changes, and manually maintained filters to block repeated bad requests. Toward the same goal, control systems today are supposedly disconnected from external networks (especially when attacks are suspected), although not consistently. Embedded systems typically have no real protection for survivability from malicious attacks (apart from some physical security), even when external connections exist.

The current metrics for survivability, availability, and reliability of time-critical systems are based on the probabilities of natural and random failures (e.g., MTBF). These metrics typically ignore intentional attacks, cascading failures, and other correlated causes or effects. For example, coordinated attacks and insider attacks are not addressed in most current approaches to survivability. One often-cited reason is that we do not have many real-world examples of intentional well-planned attacks against time-critical systems. However, because of the criticality of the systems considered here and because of many confirmed vulnerabilities in such systems, we cannot afford to wait for such data to be gathered and analyzed.

What is the status of current research?

The current state of research can be partitioned into three areas: understanding the mission and risks; survivability architectures, methods, and tools; and test and evaluation.

Understanding the Mission and Risks. We need to better understand the time-critical nature of our systems and their missions. We also need to better understand the risks to our systems with respect to impaired survivability. The concept of risk typically includes threats, vulnerabilities, and consequences. (Experiences with the design and operation of critical infrastructure systems would be helpful toward these goals.) Some methodologies and tools exist in this area, but many risk analysis methods are imprecise and suffer from limited data for one or several parameters. However, the recent efforts by Haimes et al. and Kertzner et al. are worth noting [Hai+2007, Ker+2008].

Survivability Architectures, Methods, and Tools. Efforts in this area include the large body of work in fault tolerance for systems and networks (e.g., see [Neu2000] for many references). A previous major R&D program in this area was DARPA's OASIS (Organically Assured and Survivable Information Systems), documented in the Third DARPA Information Survivability Conference and Exhibition [DIS2003]. Some work in the area of survivable control systems is also under way in the I3P program (www.thei3p.org/research/srpcs.html). However, considerable effort is needed to extend fault tolerance concepts to survivability (including intrusion tolerance) and to pursue automated and coordinated attack response and recovery.

Test and Evaluation. We need to be able to test and evaluate the time-critical elements of systems. Some testbed efforts have made general network testing infrastructures available to researchers (for example PlanetLab, ORBIT, and DETER). Some other existing testbeds are available only to restricted groups, such as military or other government research laboratories. However, testing of survivability is inherently unsatisfactory, because of the wide variety of adversities and attacks, some of which may arise despite being highly improbable. In addition, testbeds tend to lack realism.

FUTURE DIRECTIONS

On what categories can we subdivide the topics?

This topic is divided into three categories, as suggested in the preceding section: **understanding the mission and risks; survivability architectures, methods, and tools;** and **test and evaluation.**

Survivability architectures, methods, and tools are further divided into **protect**, **detect**, and **react** subcategories. Table 7.1 provides a summary of the potential approaches.

What are the major research gaps?

As an attribute of trustworthiness, survivability depends on trustworthy computer systems and communications and trustworthy operations relating to security, reliability, real-time performance where essential, and much more. Thus, it is in essence a meta-requirement. Its dependence on other subrequirements must be made explicit. (For example, see [Neu2000].) The absence of meaningful requirements for survivability is a serious gap in practice and is reflected in various gaps in research—for example, the inability to specify requirements in adequate detail and completeness, and the inability to determine whether specifications and systems actually satisfy those requirements.

Understanding the Mission and Risks

- Rigorous definitions of properties and requirements are needed that can apply in a wide range of application environments. These include concepts such as response time, outage time, and recovery time. Specific sets of requirements will apply to specific systems. We need processes and methods to identify and locate time criticality in systems and to express them in a rigorous manner. Similarly, we need to be able to identify and quantify consequences, which could be life-critical, environmental, or financial. The interaction between physical and digital systems needs to be understood with greater fidelity.
- Interdependencies among systems and infrastructures need to be analyzed. We need to understand the extent to which a survivability failure in one system can cause a failure in another system, and the ways in which survivability properties can compose.
- We need to be able to build models of systems, threats, vulnerabilities, and attack methods. These models should include evolution of attacks and blended threats that combine independent and correlated attack methods.
- There is no one-size-fits-all architecture. Some systems will be embedded and centralized; some will be networked and distributed. However, composable, scalable trustworthy systems (Section 1) are likely to play a major role.

Survivability Architectures, Methods, and Tools

Protect (protection that does not involve human interaction)

- We need families of architectures with scalable and composable components that can satisfy critical trustworthiness requirements for real-time system behavior. We need to understand how to balance confidentiality and integrity against timely availability. Traditional security mechanisms tend to either introduce human timescales or latency on a machine timescale and could thereby impair availability. Techniques for protecting integrity could

TABLE 7.1: Potential Approaches

Category	Definition	Potential Approaches
Protect	Protect systems from all relevant adversities in the system's environment.	Inherently survivable system architectures with pervasive requirements.
Detect	Detect potential failures and attacks as early as possible.	Broadly based adversity detection systems that integrate misuse detection, network monitoring, etc.
React	Remediate detected adversities and recover as extensively as possible.	Use situational awareness and related diagnostics to assess damage; anticipate potential recovery modes.

improve survivability, but not necessarily. Some integrity protection mechanisms, such as checksums, could introduce vulnerabilities if the checksums could be manipulated or made unavailable. Better techniques are needed to ensure self-monitoring and self-healing system capabilities, as well as autonomous operation. Distributed systems must also be considered, not just embedded systems. Trustworthy management (including control, security, and integrity), timely delivery of distributed data, and heterogeneous sensors will be particularly important. Survivability also requires protection against attacks, insider misuse, hardware faults, and other adversities. It may also need to limit dependence on untrustworthy components, such as complex operating systems that need frequent patches. Above all, operational interfaces to human controllers will be vital, especially in emergency situations.

- We need new communication protocols that are designed for survivability. For example, a protocol could make sure that an attacker needs to spend more resources than the system needs to expend to defend itself while preserving its time-critical properties. Frequency hopping and SYN cookies are examples of approaches using this principle. Extending or replacing TCP/IP, Modbus, and other protocols might be considered.
- We need to understand how core functions of systems can be isolated from functions that can be attacked, so that the time-critical properties of the core functions are preserved even when the systems are attacked. Research is needed on predictably trustworthy resource allocation and scheduling applicable to each of a wide range of different system architectures with different types of distributed control.
- We need to explore how we can achieve useful redundancy, with adequate assurance that single points of failure are not present.
- We must be able to identify and prevent the possibilities of cascading failures. In particular, we need mechanisms that detect and stop cascading failures faster than they can propagate. This is a complex problem that needs large testbeds and new simulation methodologies.
- Common-mode failures are a challenge in monocultures, whereas system maintenance is problematic in diversified and heterogeneous systems. Techniques are needed to determine appropriate balances between diversity and monoculture to achieve survivability in time-critical systems.
- Considerable effort is being devoted to developing hypervisors and virtualization. Perhaps these approaches could be applied to integrating COTS components into systems that are more survivable.
- We need substantive methods for composable survivability. See Section 1 (Scalable Trustworthy Systems) for a more detailed discussion on composability. We need tools for reasoning about composable survivability, including assurances relating to identity and provenance of components (Sections 6 and 9, respectively) and life cycle evaluations (Section 3). For example, survivability claims for a system composed of components should be derivable from survivability claims for components. Developing and deploying generic building-block platforms for composable survivability would be very useful.
- For networks, we need to explore the trade-offs between in-band and out-of-band control with respect to survivability, time criticality, and economics.
- We must be able to ensure survivability for services on which our time-critical systems depend. For example, all systems depend on some form of power source, and the survivability of the system can never be better than the survivability of its power sources. Other services to consider are cooling, communications, DNS, and GPS.
- We need to investigate functional distribution as a strategy for time-critical survivability and

consider challenges related to that strategy. Issues to investigate include the use of robust group communication schemes—peer-to-peer and multicast for time-critical systems.

- Detection and recovery mechanisms themselves (see below) need to be protected, to make sure they cannot be disabled or tricked into reaction.

Detect

To detect when the survivability of a time-critical system is at risk, we need to have sophisticated and reliable detection methods. This capability requires runtime methods to detect loss of time-critical system properties, such as degradation, and predict potential consequences. The following topics need investigation:

- Self-diagnosis (heartbeats, challenge-response, built-in monitoring of critical functions, detection of process anomalies).
- Intrinsically auditable systems (systems that are by design instrumented for detection).
- Network elements that participate and collaborate on detection.
- Human-machine interfaces that enable better detection and better visualization.
- Protocols that support closed-loop design (confirmation of actions).

React

When we have detected that survivability is at risk, we need to react to make sure that survivability is preserved. The following approaches to reaction need to be investigated:

- Self-healing systems that deploy machine-time methods to restore time-critical system properties.
- Graceful degradation of service (connection with mission understanding requirements).
- Predictable reactions with appropriate timeliness.
- Strategies for course of action when intervention is required (scenario planning before reaction is needed, cyber playbook).
- System change during operation (to break adversarial planning, to make planned attacks irrelevant).
- Coordinating reaction with supporting services (e.g., tell ISP to reconfigure routing into user's network, real-time black hole).
- Tarpitting, that is, slowing down an attacker without slowing down critical system functions.
- Bringing undamaged/repaired components back online via autonomous action (no human intervention). This includes reevaluation of component status and communication flows (routing, ad-hoc networks).

What are the challenges that must be addressed?

Significant advances in attacks on survivability may require research in new areas. Breadth of service environments can be important, but "depth" of hardening can also be important, as can affordability—an approach that is cost prohibitive will not be very widely adopted.

What R&D is evolutionary and what is more basic, higher risk, game changing?

Near term

- Realistic, comprehensive requirements
- Existing protocols
- Identification of time-critical components

Medium term

- Detection
- Strategies for reaction
- Experimentation with trustworthy protocols for networking and distributed control, out-of-band signaling, robustness, and emergency recovery
- Higher-speed intercommunications and coordination
- Development tools
- System models

Long term

- Evaluatable metrics
- Establishment of trustworthy protocols for networking and distributed control
- Self-diagnosis and self-repair
- Provisioning for automated reaction and recovery

Resources

Making progress on the entire set of in-scope systems requires focused

research efforts for each of the underlying technologies and each type of critical system, together with a research-coordinating function that can discern and understand both the common and the disparate types of solutions developed by those working on specific systems. An important role for the coordinating function is to expedite the flow of ideas and understanding among the focused groups.

For a subject this broad and all-encompassing (it depends on security, reliability, situational awareness and attack attribution, metrics, usability, life cycle evaluation, combating malware and insider misuse, and many other aspects), it seems wise to be prepared to launch multiple efforts targeting this topic area.

Measures of success

Success should be measured by the range of environments over which the system is capable of delivering adequate service for top-priority tasks. These environments will vary by topology and spatial distribution: number, type, and location of compromised machines; and a broad range of disruption strategies.

What needs to be in place for test and evaluation?

Many issues are relevant here:

- **Metrics for survivability:** determining which existing metrics (MTBF, etc.) are applicable, which measures of success are appropriate, what additional aspects of survivability and time criticality should be measured (not covered by existing metrics). Resilience must be possible in the face of unexpected inputs, when some partial degree of service must still be provided, with appropriate recovery time. Attack efforts in testing need to be appropriately high.
- **Measuring** the relationships between complexity and time criticality is desired, especially when a system requires faster-than-human reactions.
- **High-fidelity simulations**, including: how to simulate physical aspects together with control functions, integrate security in testing and simulation, and validate the simulation. Appropriate degrees of fidelity, and determining that a simulation is sufficiently realistic.
- **Private industry** needs to be engaged.
- **Analytical models** should be developed based on simulations.
- **Red Teaming** to assess structured survivability, with red teams employing domain-specific skills.
- **Adversarial modeling** that seeks to understand the threat to time-critical systems.

To what extent can we test real systems?

- **Testing of large systems:** survivability is not easy to test in a very large and complex system, such as an electric power grid. Relevant issues include: how to share access to existing testbeds and how to compose results of subsystem tests.
- **Research infrastructures** that are needed to support research in this area include a "library" of devices: keep a copy of every reasonably sized and priced manufactured device (compare this with seed banks). Also, keep templates or models of devices for use in design and evaluation.
- **Access to real-world normal and attack data** and system designs for evaluating research results is needed for all types of systems covered in this section, not just for typical data but also for extreme cases. Issues concerning proprietary data and data sanitization need to be addressed, including post-incident data and analysis such as flight data records; and integration of testbeds (wireless, SCADA, general IT), enabling testbed capabilities to be combined.

References

[Avi+2004] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1):11-33, January-March 2004.

[DIS2003] 3rd DARPA Information Survivability Conference and Exposition (DISCEX-III 2003), 22-24 April 2003, Washington, DC, USA. IEEE Computer Society 2003, ISBN 0-7695-1897-4.

[Ell+1999] R.J. Ellison, D.A. Fisher, R.C. Linger, H.F. Lipson, T. Longstaff, and N.R. Mead. *Survivable Network Systems: An Emerging Discipline*. Technical Report CMU/SEI-97-TR-013, Carnegie Mellon University, May 1999.

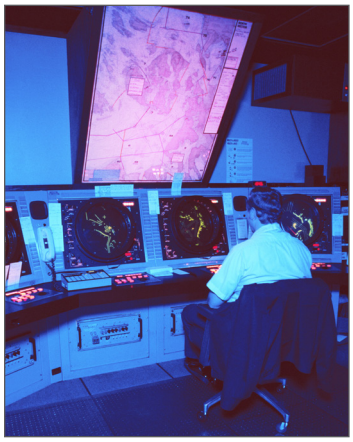
[Hai+2007] Yacov Y. Haimes, Joost R. Santos, Kenneth G. Crowther, Matthew H. Henry, Chenyang Lian, and Zhenyu Yan. Analysis of Interdependencies and Risk in Oil & Gas Infrastructure Systems. I3P Research Report No. 11, June 2007 (http://www.thei3p.org/docs/publications/researchreport11.pdf).

[Ker+2008] Peter Kertzner, Jim Watters, Deborah Bodeau, and Adam Hahn. *Process Control System Security Technical Risk Assessment Methodology & Technical Implementation*. I3P Research Report No. 13, March 2008 (http://www.thei3p.org/docs/publications/ResearchReport13.pdf).

[Neu2000] P.G. Neumann. Practical Architectures for Survivable Systems and Networks. SRI International, Menlo Park, California, June 2000 (http://www.csl.sri.com/neumann/survivability.html).

8. Situational Understanding and Attack Attribution

BACKGROUND



What is the problem being addressed?

Situational understanding is information scaled to one's level and areas of interest. It encompasses one's role, environment, the adversary, mission, resource status, what is permissible to view, and which authorities are relevant. The challenges lie in the path from massive data to information to understanding, allowing for appropriate sharing at each point in the path.

The questions to be answered, in rough order of ascending difficulty, are the following:

- Is there an attack or misuse to be addressed (detection, threat assessments)?
- What is the attack (identification, not just IDS signature)?
- Who is the attacker (accurate attribution)?
- What is the attacker's intent (with respect to the present attack as well as predicting behavior over time)?
- What is the likely impact?
- How do we defend (autonomous enterprises and the community as a whole)?
- What (possibly rogue) infrastructure enables the attack?
- How can we prevent, deter, and/or mitigate future similar occurrences?

Situational understanding includes the state of one's own system from a defensive posture irrespective of whether an attack is taking place. It is critical to understand system performance and behavior during non-attack periods, in that some attack indicators may be observable only as deviations from "normal behavior." This understanding also must include performance of systems under stress that are not caused by attacks, such as a dramatic increase in normal traffic due to sudden popularity of a particular resource.

Situational understanding also encompasses both the defender and the adversary. The defender must have adversary models in order to predict adversary courses of action based on the current defensive posture. The defender's system-level goals are to deter unwanted adversary actions (e.g., attacking our information systems) and induce preferred courses of action (e.g., working on socially useful projects as opposed to developing crimeware, or redirecting attacks to a honeynet).

Attack attribution is defined as determining the identity or location of an attacker or an attacker's intermediary. Attribution includes the identification of intermediaries, although an intermediary may or may not be a willing participant in an

attack. Accurate attribution supports improved situational understanding and is therefore a key element of research in this area. Appropriate attribution may often be possible only incrementally, as situational understanding becomes clearer through interpretation of available information.

Situational understanding is larger than one user, or possibly even larger than one administrative domain, and addresses what is happening through consideration of a particular area of interest at a granularity that is appropriate to the administrator(s) or analyst(s). In particular, situational understanding of events within infrastructures spanning multiple domains may require significant coordination and collaboration on multiple fronts, such as decisions about when/whether to share data, how to depict the situation as understanding changes over time, and how to interpret or respond to the information. Attribution is a key element of this process, since it is concerned with who is doing what and what should be done in response.

What are the potential threats?

Situational understanding addresses a broad range of cyber attacks, specifically including large-scale and distributed attacks, where it is felt that adversary capabilities are outstripping our ability to defend critical systems. Inability to attribute sophisticated attacks to the original perpetrator leads to a growing asymmetry in cyber conflict.

In this topic area, we are concerned chiefly with the universe of cyber attacks within the information systems domain and how our decision makers interpret, react to, and mitigate those attacks. Of special concern are attacks on information systems with potentially significant strategic impact, such as wide-scale power blackouts or loss of confidence in the banking system. Attacks may come from insiders, from adversaries using false credentials, from botnets, or from other sources or a blend of sources. Understanding the attack is essential for defense, remediation, attribution to the true adversary or instigator, hardening of systems against similar future attacks, and deterring future attacks. Attribution should also encompass shell companies, such as rogue domain resellers whose business model is to provide an enabling infrastructure for malfeasance. There are numerous areas of open research when it comes to these larger questions of attribution. For example, we have not adequately addressed digital fingerprinting of rogue providers of hosting services. (See also Section 9.)

There have been numerous widely publicized large-scale attacks launched for a variety of purposes, but recently there is a consensus that skilled nonstate actors are now primarily going after financial gain [GAO2007, Fra2007]. Click fraud, stock "pump and dump," and other manipulations of real-time markets prove that it is possible to profit from cybercrime without actually taking down the systems that are attacked. In this context, situational understanding should clearly encompass law enforcement threat models and priorities, as well as how financial gains can accrue.

For state actors, the current concern is targeting of our critical infrastructures and key government systems. Adversaries may be able to exfiltrate sensitive data over periods of time, again without actually taking down the targeted systems. Here, situational understanding should clearly include understanding of government threat models and concerns. Sharing such understanding is particularly important—and sensitive in the sense that it is likely to lead to recognition of additional weaknesses and vulnerabilities.

In addition, the more serious attacks now occur at two vastly different timescales. The classic fear is cyber attacks that occur faster than human response times. Those attacks are still of concern. However, another concern is "low and slow" and possibly stealthy attacks that break the attack sequences into a series of small steps spread over a long time period. Achieving situational awareness for these two ends of the continuum is likely to require very different approaches.

Who are the potential beneficiaries? What are their respective needs?

Although all computer users and all consumers of information systems products are potential victims of the broad range of attacks we address, and would benefit from improved situational awareness, we are primarily seeking tools and techniques to help the communities whose challenges and needs are given in Table 8.1—although this is not an comprehensive set.

Because of time criticality for responding to certain cyber attacks, and hence the need to tie these to situational awareness, we consider developers and users

TABLE 8.1: Beneficiaries, Challenges, and Needs

Beneficiaries	Challenges	Needs
System Administrators	Overwhelmed by attacks buried in massive data volumes. Limited visibility beyond own domain.	Timely detection, presentation, sharing with peers across administrative boundaries. Effective remediation.
Service Providers	Service continuity in spite of large-scale attacks. Understanding emerging attacks. Sharing with peers.	Attack attribution. Identify and quarantine compromised systems. Reliable IP mapping to jurisdiction to support effective cooperation with law enforcement.
Law Enforcement	Identify and prosecute perpetrators (individuals and emerging cybercrime syndicates).	Coordination with service providers and administrators. Data collection, presentation, and analysis of forensic quality. Attribution to ultimate perpetrator.
Civil Government	Continuity in spite of large-scale attacks on government and civilian systems. Coordination of national-level response.	Detection of attacks. Early identification of attacks on critical infrastructure sectors. Sharing with private sector as well as state/local agencies. Attribution.
Military	Prevent attacks on defense systems. Maintain system continuity in spite of attacks. Prevent exfiltration of critical data.	Early detection and identification of attacks. Attribution. All of the above.

of autonomic response systems as part of the customer base for advances in this topic area.

What is the current state of the practice?

Situational understanding currently is addressed within administrative domains through intrusion detection/prevention systems and security event correlation systems, with much of the analysis still done through manual perusal of log files. There have been efforts to provide visualizations and other analytical tools to improve the ability to comprehend large amounts of data. These are largely special purpose and found within research laboratories rather than being used widely within the field. Sharing security-relevant information across domains is essential for large-scale situational understanding but is currently accomplished via ad hoc and informal relationships. In a few instances, data is shared across organizations, but normally the kinds of information shared are limited (e.g., only network packet headers).

Intrusion detection/prevention technology is widely deployed, but many question how much longer it will be effective as traffic volumes grow, attacks get more subtle, signature bases grow correspondingly larger and unable to cope with new attacks, and attackers use encryption, which makes packet payload signature analysis difficult. Response to large-scale attacks remains to a large degree informal, via personal trust relationships and telephone communications. This situation makes it difficult or impossible to achieve very rapid response or cooperation between domains where the administrators do not know and trust each other. (For example, how can an administrator in Domain A prove that a customer of Domain B is an attacker, and thereby persuade an administrator in that domain to take corrective action?)

Industry has made significant progress in the area of event/data correlation, with several security information and event management (SIEM) commercial products widely deployed in the market. These offer considerable value in timely data reduction and alarm management. However, with respect to visualization and presentation on a massive data scale, these systems are inadequate and do not have scope well beyond organizational boundaries.

We need to consider the viewpoint of the defender (end host, infrastructure component, enterprise, Internet). An

ISP wants an "inward" view of enterprise customers since cooperative security benefits from each domain's filtering outbound attack traffic from that domain (**egress filtering**). A defender at an edge router is also looking outward at its peers to monitor the inbound flows for attack traffic (**ingress filtering**). This ingress filtering is essential to the cooperative awareness and response mentioned above.

Lack of trust between providers, issues of scalability, and issues of partial deployment of defenses make attribution difficult in many cases. Privacy regulations and the very real concern that data sanitization techniques are ineffective also present barriers. The differing legal regimes in different countries, or within different areas of governments within the same country, inhibit attribution as well. There is a need for international dialogue in how to handle cybersecurity incidents, so that attackers can either be identified and prosecuted or otherwise deterred from future wrongdoing.

Progress is being made in many areas important to situational understanding, including attribution. Protocols such as IPsec and IPv6's extension headers for authentication may improve the situation in the sense that spoofing the attack source is more difficult than in current IPv4 networks. However, these message authentication techniques do not solve the underlying problem of compromised machines being used to attack third parties. Thus, there is an important linkage between this topic and the topic addressing malware (see Section 5).

There are several forums for security event information sharing, such as SANS Internet Storm Center's dshield [ISC], which describes itself as a cooperative network security community, and PhishTank [Phi], which allows the defender community to contribute known or suspected instances of phishing attacks. Phishing refers to a broad class of fraudulent attempts to get a user to provide personal information that the phisher can subsequently use for identity theft, identity fraud, unlawful financial transactions, and other criminal activity.

For reasons ranging from customer privacy and concerns about revealing defensive posture to legal liability issues, only limited meaningful progress has been made in the area of interdomain security information sharing and in determining attacker location and intent.

What is the status of current research?

Research in attack detection has continued along the path of faster signature development and propagation, seeking to reduce the time window in which zero-day attacks have an impact.

Egress filtering is increasingly used to identify internal assets that may be currently compromised. This egress filtering (or, more generally, "unbiased introspection") also applies to ISPs, enterprises, and home computers.

Scalable information processing (e.g., data reduction), data mining, statistical analysis, and other similar techniques are applied to situational understanding. There are significant challenges and opportunities as link speeds become faster and data storage becomes cheaper.

In the area of attribution, there is active research in traceback techniques. However, most methods depend on cooperative defense and do not function well with less than universal deployment. Skilled attackers easily evade most currently deployed traceback systems.

There has been some research in attacker intent modeling, with an objective to predict the attacker's next steps, but this has had only limited success. In addition, most academic research in the cybersecurity field uses inadequate adversary models that do not capture how high-level adversaries actually attack complex systems. As mentioned previously, the short-term goal is modeling adversary behavior to generate better attack indicators. The long-term goal is to deter unwanted behaviors and to promote appropriate behaviors (e.g., working for a legitimate organization as opposed to organized crime) via improved attribution. Most research in this area is emphasizing the short-term goal rather than the longer-term goal.

Sharing actionable data while respecting privacy, authenticating the shared information in the absence of interdomain trust, the economy of sharing (sharing marketplace), and sharing with privacy and anonymity are important research issues (see Section 10). Policy and legal barriers to sharing also need to be addressed, in addition to the difficult

technical questions. Sharing lets one know if he or she is part of an attack and needs to take action, and also lets one see the global picture. Some of the legal framework from the PREDICT data repository may be applicable (http://www.predict.org). There are also examples in international scientific collaborations involving information systems that could be considered in ways to collectively identify threats. Another model for sharing is seen in the international honeynet community.

There are different variants of attribution. In closed user communities, users often consent to monitoring as a condition for system access, so it is easier to assert who did what. A consent-to-monitoring policy is not likely to be implemented globally, so attribution of attacks that come in from the Internet will remain difficult. This second type of attribution should be balanced against the need for anonymity and free speech concerns arising from requiring that all traffic to be subject to attribution.

FUTURE DIRECTIONS

On what categories can we subdivide this topic?

We frame this topic area on the following categories:

- **Collection.** Identify what data to collect; develop methods for data collection, preservation of chain of custody (see Section 9), validation, and organization.
- **Storage.** Decide how to protect data in situ, efficiently access stored data, and establish reporting responsibilities, assure integrity, and how long to store data and in what form.
- **Analysis.** Analyze the collected data to abstract out meaning, potentially seek additional information for consolidation, identify security incidents and compute relevant metadata.
- **Presentation.** Distill security incidents and related contextual information to form enterprise-level situational awareness; enable responses while maintaining forensic quality for attribution. Presentation may involve data sanitization or modification to comply with privacy or classification-level requirements on who is allowed to view what.
- **Sharing.** Develop sharing awareness across independent domains and mechanisms to present relevant data to appropriate communities, such as network operators and law enforcement, and preserve privacy of users, sensitive corporate and national-security data, and system defensive posture.
- **Reaction.** Determine local and cross-domain course of action to mitigate events. This includes measures to stop further damage, fix damage that has occurred, proactively change security configurations, and collect forensics to enable attribution and prosecution.

This framework may be considered an adaptation of John Boyd's OODA loop (Observe, Orient, Decide, Act (http://en.wikipedia.org/wiki/OODA_Loop). By analogy to physical security systems, "reaction" might be further broken out into delay, response, and mitigation steps. Some courses of action by the defender might delay the adversary from achieving the ultimate objective of the attack. This buys time for the defender to mount an effective response that thwarts the adversary's goal. Another response might be to seek out additional information that will improve situational awareness. If an effective response is not possible, then mitigation of the consequences of the adversary's action is also a valuable course of action. Many responses may require coordination across organizational boundaries, and shared situational awareness will be important in supporting such activities.

What are the major gaps?

Attack signature generation and propagation are falling short, as many "legacy attacks" are still active on the Internet years after they were launched. Legacy attacks persist for many reasons, such as poor system administrative practices or lack of support for system administration, a proliferation of consumer systems not under professional system administration but with a high-bandwidth connection, reemergence of older machines after being in storage without appropriate attention (e.g., travel laptops put back into service), or use of legacy code or hardware in new applications or devices. This persistence indicates that research is needed into better tools for system administration, as well as for survivability of well-administered systems in an environment where many other systems are

poorly maintained. Also, the ability to quickly scrutinize new applications and devices to see whether legacy flaws have been reintroduced would be beneficial.

There remain significant gaps in the intrusion detection field, and currently deployed intrusion detection systems (IDS) fall short of needs, especially with respect to enabling distributed correlation. In particular, approaches that include ever-growing signature sets in attempting to identify the increasing variety of attacks may be approaching the end of their usefulness; alternative approaches are clearly needed.

Detection of attacks within encrypted payloads will present an increasingly serious challenge. Many botnets now use encrypted command and control channels. There are researchers investigating techniques that take advantage of this, such as using the presence of ciphertext on certain communications channels as an attack indicator. However, it is likely that the fraction of encrypted traffic will increase under legitimate applications, and thus alternative approaches are once again needed.

Attribution remains a hard problem. In most modern situations, it is useful to get as close as possible to the ultimate origin (node, process, or human actor). However, doing so touches on privacy, legal, and forensic issues. For example, public safety argues for full attribution of all actions on the Internet, while free-speech rights in a civil society are likely to require some forms of socially approved anonymity. We also need to define the granularity of attack attribution. In this respect, attribution could apply within a single computer or local network, but it could also be sufficient to provide attribution within a domain, or even a country. Moreover, adversaries are getting better at hiding the true origin of their attacks behind networks of compromised machines (e.g., botnets), and throwaway computers may become as common as throwaway cell phones as prices drop. Adversaries increasingly use techniques such as fast flux, where the DNS is rapidly manipulated to make identification and takedown of adversary networks difficult [Hol2008].

What are some exemplary problems for R&D on this topic?

Collect and Store Relevant Data. Understand how to identify, collect, and ultimately store data appropriate to the form of situational awareness desired. This might involve network-centric data such as connectivity with peers over time, archives of name resolution, and route changes. In addition, data may need to be combined and/or sanitized to make it suitable for sharing or downstream retrieval, such as with lower-layer alerts, local as well as external view, system- and application-level alerts, packet contents supporting deep packet inspection on demand without violating privacy or organizational security, archives to support snapshots and history, and externally deployed monitoring infrastructure such as honeynets. Finally, data outside networks and hosts is also relevant, such as "people layer" knowledge, as in tracking the so-called Russian Business Network (RBN) over time.

In addition to the database hurdles (such as scale and organization) that must be overcome in the collection of these diverse sources, it is in the interest of the adversary to poison these data sources. Research is needed so that data provenance can be maintained. (See Section 9.)

Analysis on Massive Data Scales. The analysis or evaluation process must consider the massive scale and heterogeneity of data sources and the fact that most of the data arriving from the above sources is uninteresting chaff. The data and analysis should support a variety of granularities, such as Border Gateway Protocol (BGP) routes, DNS queries, domains in country-code top level domains (TLDs), repeated patterns of interaction that arise over the course of months or years, and unexpected connections between companies and individuals. These derived quantities should themselves be archived or, alternatively, be able to be easily reconstructed. The availability of these data sources plays an important role in enabling attack attribution and also contributes to an incremental building of situational awareness.

Novel Approaches to Presentation in Large-Scale Data. The massive scale of the data poses challenges to timely, compact, and informative presentation. Scalable visualization, visualization with accurate geolocation, and zoomable visualization at varying levels of detail are just some of the difficult problems. Maintaining the ability to delve into the original data as well as broaden out to a high-level, people-aware view is an area for future research.

Collaborative Collection, Vetting, and Archiving. Collaborative collection of non-open data and the subsequent vetting, archiving, correlation (for example inferring routes collaboratively), and generation of useful metadata are important research issues. Numerous database issues arise, including processing of huge repositories, definition and derivation of meaningful metadata such as provenance, validation of inputs, and multilevel security (MLS). Such an archive would support both research and operations. There are serious questions as to what to share and to what degree, and these questions may occur at multiple levels. Examples include what one controls, what one shares in a trusted community, and what we can observe about an uncooperative and possibly adversarial entity.

Cross-Boundary Sharing of Situational Understanding. Crossing organizational boundaries may require reputation systems and other ways of quickly determining when it might be safe to share information that cannot itself be gamed. It may be possible to leverage research in reputation in peer-to-peer (P2P) systems. Multiple issues arise with modern approaches. Sparse reports may be misleading, because voting mechanisms may not allow determining truth. Proving that only one organization is under attack may be difficult (likely to require submitting traffic samples that may reveal defensive posture, and subject to the possibility of spoofing). We require research in enabling technologies to promote sharing across organizational boundaries.

Situational Understanding at Multiple Timescales. We must be aware that there are multiple timescales at which situational understanding must be inferred and presented. For low and slow attacks, such as those involved in insider-threat investigations, the attack traffic may occur over long time spans (years or decades) and encompass multiple ingress points. In contrast, autonomic response requires millisecond situational understanding. For the human consumer, the timescale is intermediate.

Some exemplary approaches are summarized in Table 8.2.

What R&D is evolutionary and what is more basic, higher risk, game changing?

Along the collection dimension, near- and medium-term areas include identification of data types, sources, collection methods, and categorization; directed data selection; and instrumentation of software and hardware components and subsystems. Long-term research may consider systems that are intrinsically enabled for monitoring and auditing. Challenges include the rapid growth of data and data rates, changing ideas about what can potentially be monitored, and privacy issues. (See Section 10.)

With respect to analysis, there is consensus that the current signature-based approaches will not keep up with the problem much longer, because of issues of scale as well as system and attack complexity. Attack triage methods should be examined in the short term. Traffic encryption and IPv6 may render many attack vectors harder but may also make analysis more difficult. In the long term, conceptual breakthroughs are required

TABLE 8.2: Exemplary Approaches

Category	Definition	Sample Solutions
Collect and Analyze Data	Understanding threats to overall trustworthiness and potential risks of failures	Broad-based threat and misuse detection integrating misuses and survivability threats
Massive-Scale Analysis	New approaches to distributed system and enterprise attack analysis	Trustworthy systems with integrated analysis tools
Situational Understanding across boundaries and multiple timescales	Interpretation of multiple analyses over space and time	Intelligent correlated interpretation of likely consequences and risks

to stay even with or ahead of the threat. For example, some botnet command and control (C2) traffic is already on encrypted channels. Ideally, intrinsically monitorable systems would permit an adversary little or no space to operate without detection, or at least permit observation that could be turned to detection with additional analysis. Such systems detect attacks without a signature base that grows essentially without bound. They also permit one to reliably assert that a system is operating in an acceptably safe mode from a security standpoint. Additional approaches are needed that address monitoring and analysis in system design.

The state of the art tends to rely on detection. Some limited progress has been made to date on predicting attackers' next steps or inferring attacker intent. Advances in target analysis will better identify what is public and thus presumed known to the adversary. This work may lead to solutions whereby defenders manipulate the exposed "attack surface" to elicit or thwart attacker intent, or use cost-effective defenses that increase protections when it is predicted they are needed. Correlated attack modeling advances are appropriate to pursue as a medium-term area. Game theoretic and threat model approaches have made limited headway in this field but should be considered as long-term research. Threat and adversary modeling may also support advances toward attribution and the ultimate goal of deterring future cyber attacks. This is suitable for medium- to long-term research.

Information presentation will require continued advancements in data reduction, alarm management, and drill-down capability. In the near term, the emerging field of visual analytics may provide useful insights, with new visualization devices presenting opportunities for new ways of viewing items. An emerging challenge in displaying situational awareness is the increase in reliance both on very large (wall-size) viewing screens and on very small handheld screens (e.g., BlackBerries). A suggested long-term effort is to consider alternative metaphors suited to the various extremes available, including such options as the scrollable, zoomable map. Inference and forecasting are also appropriate for long-term efforts. We should build on the research in information presentation for human understanding and response. Another hard problem is visualization of low and slow attacks. Near- and medium-term research is needed in how to assess the way different situational awareness presentation approaches affect an analyst's or administrator's ability to perform.

Presentation approaches need awareness as to whether the consumer is a human or an autonomous agent; reliance on intelligent agents or other forms of automated response means that these elements will also need "situational awareness" to provide context for their programmed behaviors. We require research to enable agent-based defenses in instances where action is needed at faster than human response times. This is a presentation issue that ought to be addressed in the medium term, and a sharing issue when agent-to-agent cooperation is required in the long term. It is important to keep in mind that autonomous response may be an attack vector for the adversary, and the ability to change the situational awareness information presented to agents or other autonomous response vehicles is a potential vulnerability.

Sharing relevant information spans the gamut of levels from security alerts to sharing situational understanding obtained from analysis. Sharing can enable global situational understanding and awareness, support reliable attribution, and guide local response appropriate to the global picture. Research is needed to determine how to achieve sharing with adequate privacy protections, and within regulatory boundaries, what to share across autonomous systems, and possible market mechanisms for sharing. The issue of liability for misuse or for fraudulent or erroneous shared data will need to be addressed.

Research in appropriate reaction has both local (within an enterprise, within an autonomous systems) and global (across enterprise and autonomous systems) components. Ideally, the output of current and previous research results should support an effective course of action. When this is shared between entities, the shared information should support effective local reaction, while preserving privacy along with other information sanitization needs. Research is required, for example in authenticating the authors of actionable information and proof that a recommended course of action is appropriate. Research is also required in alternatives to malfeasor blocking (it may be preferable to divert and observe), remediation of compromised assets (a need also present in the malware research topic), and exoneration in the case of false

positives. Although response and reaction are not directly a part of situational understanding, situational understanding is needed to enable response and reaction, and situational understanding may drive certain kinds of responses (e.g., changing information collection to improve attribution). Thus, advances in reaction and response techniques directly affect the kind of situational awareness that is required.

Resources

Situational understanding requires collection or derivation of relevant data on a diverse set of attributes. Some of the attributes that support global situational understanding and attack attribution are discussed above relating to the kinds of data to collect. A legal and policy framework, including international coordination, is necessary to enable the collection and permit the exchange of much of this information, since it often requires crossing international boundaries. In addition, coordination across sectors may be needed in terms of what information can be shared and how to gather it in a timely way. Consider an attack that involves patient data information systems within a hospital in the United States, a military base in Germany, and an educational institution in France. All three institutions have different requirements for what can and cannot be shared or recorded.

Modifications to U.S. law and policy may be needed to facilitate data sharing and attack attribution research. As an example, institutional review boards (IRBs) play an important role in protecting individuals and organizations from the side effects of experimentation that involves human subjects. In many cases, the IRBs are inadequately equipped to handle cybersecurity experiments, which are crucial to understanding attackers' intent and next steps. Government could play a role in ensuring that IRBs are better equipped to expedite attack attribution research. A set of best practices would be beneficial in this area.

Government roles also include developing policy, funding research (complementing industry), and exerting market leverage through its acquisition processes. There is government-sponsored research in intrusion detection, software engineering for security, malware analysis, traceback, information sharing, scalable visualization, and other areas that affect this topic. Government has also implemented fusion centers, common databases for experimentation, and testbeds, supporting collaboration. Continuing these investments is crucial, particularly in the long-term range for areas that are not conducive to short-term industry investment.

This topic is particularly dependent on public-private partnerships, and the definition of the nature of these partnerships is essential. To a degree, this depends on competing visions of success. One may consider a centralized network operations center (NOC) staffed by government, industry, and researchers with a policy and procedural framework designed to allow seamless cooperation. An alternative view is a distributed capability in which different network operators share situational understanding but different parts of the picture are relevant to different system missions.

This section focuses on protection against cyber attack in the information domain. However, adversaries may choose to interleave their cyber-attack steps with attack steps in the other three domains of conflict—namely the physical, cognitive, and social domains. Research on situational understanding and attribution tools that integrate attack indicators from all four domains of conflict is also needed.

Measures of success

We will measure progress in numerous ways, such as decreased personnel hours required to obtain effective situational understanding; increased coverage of the attack space; based on mission impact, improved ability to triage the serious attacks from the less important and the ones where immediate reaction is needed from those where an alternative approach is acceptable; improved response and remediation time; and timely attribution with sound forensics. These all require reliable collection of data on the diverse set of attributes listed previously.

On the basis of these attributes, we could define measures of success at a high level within a given organization's stated security goals. For example, an organization aimed primarily at maintaining customer access to a particular service might measure success by observing and tracking over time such variables as the estimated number of hosts capable of serving information over some service, and the estimated near-steady-state number or growth trend of these machines.

Success depends on timely identification

of adversaries, propagation of defenses, and remediation of affected systems. Another measure for success is tied to a variation of the false-positive/true-positive discussion, in that effective situational understanding should allow us to accurately categorize the potential impact of a detected attack. For either actual attacks or emulated attacks on a realistic testbed, we would hope to be able to answer the following questions:

- Can we differentiate between nuisance and serious strategic attacks, for example, by identifying a targeted attack against a critical sector?
- Can we share information across informational boundaries to enable cooperative response?
- Can we quickly quarantine intermediate attack platforms?
- Can we maintain or quickly restore critical functions, perhaps according to some contingency policy of acceptable degradation?
- Can we collect actionable data for ultimate attribution?

We require a methodology to quantify mission impact. Many stakeholders have a primary need to maintain continuity of operations in spite of a large-scale attack.

What needs to be in place for test and evaluation?

Several research testbeds are online (e.g., the existing DETER lab testbed, http://www.deterlab.net) or planned; research in situational understanding would be advanced via federation of these and other testbeds to emulate scale and cross-domain issues. Large-scale simulation may provide initial rough estimates of the efficacy of particular approaches. In terms of Internet-scale situational understanding, these testbeds can support advances in the malware and botnets topic area as well.

To what extent can we test real systems?

There are test environments that allow deployment of prototype cybersecurity modules. We should consider developing an open-source framework with defined standards and interfaces, and developing relationships with entities that could deploy it. Many results from this topic require distributed deployment for meaningful test and evaluation. The honeynet community may be a good deployment platform with less resistance than commercial systems and less concern about privacy issues. Significant barriers exist in both the technical and organizational/policy domains, associated with the difficulty of protecting the privacy and security of the real systems being scrutinized.

Technologies resulting from research in this topic area range from individual-host-level components (for example, inherently monitorable systems) to global components (mechanisms for reliable geolocation). In the former category, R&D should be conducted from the start with system developers to ensure adoptability of resulting solutions. Success in the latter category may require some new frameworks in law, policy, and Internet governance.

References

[Fra2007] J. Franklin, V. Paxson, A. Perrig, and S. Savage. An inquiry into the nature and causes of the wealth of Internet miscreants. *Proceedings of ACM Computer and Communications Security Conference*, pp. 375-388, October 2007.

[GAO2007] *CYBERCRIME: Public and Private Entities Face Challenges in Addressing Cyber Threats*. Report GAO-07-705, U.S. Government Accountability Office, Washington, D.C., July 2007.

[Hol2008] T. Holz, C. Gorecki, K. Rieck, and F. Freiling. Measuring and detecting fast-flux service networks. In *Proceedings of the 15th Annual Network & Distributed System Security (NDSS) Symposium,* February 2008.

[ICA2008] *Draft Initial Report of the GNSO Fast Flux Hosting Working Group*. ICANN. December 8, 2008 (https://st.icann.org/pdp-wg-ff/index.cgi?fast_flux_pdp_wg).

[ISC] Internet Storm Center: http://www.dshield.org/about.html.

[Phi] PhishTank: http://www.phishtank.com.

9. Provenance

BACKGROUND

What is the problem being addressed?

Individuals and organizations routinely work with, and make decisions based on, data that may have originated from many different sources and also may have been processed, transformed, interpreted, and aggregated by numerous entities between the original sources and the consumers. Without good knowledge about the sources and intermediate processors of the data, it can be difficult to assess the data's trustworthiness and reliability, and hence its real value to the decision-making processes in which it is used.

Provenance refers to the chain of successive custody—including sources and operations—of computer-related resources such as hardware, software, documents, databases, data, and other entities. Provenance includes **pedigree**, which relates to the total directed graph of historical dependencies. It also includes **tracking**, which refers to the maintenance of distribution and usage information that enables determination of where resources went and how they may have been used.

Provenance is also concerned with the original sources of any subsequent changes or other treatment of information and resources throughout the life cycle of data. That information may be in any form, including software, text, spreadsheets, images, audio, video, proprietary document formats, databases, and others, as well as meta-level information about information and information transformations, including editing, other forms of markup, summarization, analysis, transformations from one medium to another, formatting, and provenance markings. Provenance is generally concerned with the integrity and reliability of the information and meta-information rather than just the information content of the document.

Provenance can also be used to follow modifications of information—for example, providing a record of how a document was derived from other sources or providing the pervasive history through successive versions (as in the Concurrent Versions System [CVS]), transformations of content (such as natural language translation and file compression), and changes of format (such as Word to PDF).

The granularity of provenance ranges from whole systems through multi-level security, file, paragraph, and line, and even to bit. For certain applications (such as access control) the provenance of a single bit may be very important. Provenance itself may require meta-provenance, that is, provenance markings on the provenance information. The level of assurance provided by information provenance systems may be graded and lead to graded responses. Note that in some cases provenance information may be more sensitive, or more highly classified, than the underlying data. The policies for handling provenance information are complex and differ for different applications and granularities.

To determine provenance accurately, we must have trustworthy systems that reliably track both usage and modification of information and other resources. As with all computer systems, security of provenance tracking cannot be absolute, and trustworthiness of provenance tracking systems will be relative to the value of the provenance to the users of the information and resources. For example, a simple change-tracking mechanism in a document preparation system may provide adequate provenance tracking from the point of view of a small group of authors collaborating in the publication of an article, even though the document change history might not be protected from unauthorized modification. On the other hand, the same mechanism may be inadequate in the context of legal discovery, precisely because the change-tracking mechanism does not guarantee the authenticity of the change history.

What are the potential threats?

Without trustworthy provenance tracking systems, there are threats to the data and to processes that rely on the data, including, for example, unattributed sources of software and hardware; unauthorized modification of data provenance; unauthorized exposure of provenance, where presumably protected; and misattribution of provenance (intentional or otherwise).

Who are the potential beneficiaries? What are their respective needs?

The legal, accounting, medical, and scientific fields are examples where provenance markings are beginning to be used. Other fields that can benefit from provenance maintenance systems include critical infrastructure providers (e.g., in SCADA and other control systems), emergency responders, military personnel, and other decision makers. Users in all these areas need reliable information obtained from many sources, communicated, aggregated, analyzed, stored, and presented by complex information processing systems. Information sources must be identified, maintained, and tracked to help users make appropriate decisions based on reliable understanding of the provenance of the data used as input to critical decisions.

In addition, new techniques are needed that will allow management of provenance for voluminous data. Part of what has made provenance easier to manage up to now is its small volume. Now, geospatial information-gathering systems are being planned that will have the capability of handling gigabytes of data per second, and the challenges of these data volumes will be exacerbated by collection via countless other sensor networks. Within 20 years, the government will hold an exabyte of potentially sensitive data. The systems for handling and establishing provenance of such volumes of information must function autonomously and efficiently with information sources at these scales.

Note that situations are likely to arise where absence of provenance is important—for example, where information that needs to be made public must not be attributable.

What is the current state of practice?

Physical provenance markings in jewelry (e.g., claiming your diamond is from a blood-free mining operation, your silver or gold is pure, and the style is not a knockoff copy of a designer's), explosive components (e.g., nitrates), and clothing have historically added value and enabled tracing of origin. Document markings such as wax seals and signatures have been used to increase assurance of authenticity of high-value documents for centuries. More recently the legal, auditing, and medical fields have begun to employ first-level authenticated provenance markings.

The current practice is rather rudimentary compared with what is needed to be able to routinely depend on provenance collection and maintenance. The financial sector (in part driven by Sarbanes-Oxley requirements) has developed techniques to enable tracking of origins, aggregations, and edits of data sets. Users of document production software may be familiar with change-tracking features that provide a form of provenance, although one that cannot necessarily be considered trustworthy.

As an example of provenance in which security of the provenance has not been a direct concern, software development teams have relied for decades on version control systems to track the history of changes to code and allow for historical versions of code to be examined and used. Similar kinds of systems are used in the scientific computing community.

What is the status of current research?

Current research appears to be driven largely by application- and domain-specific needs. Undoubtedly, these research efforts are seen as vital in their respective communities of interest.

Examples of active, ongoing research areas related to information and resource provenance include the following areas:

- **Data provenance and annotation in scientific computing.** Chimera [Fos2002] allows a user to define a workflow, consisting of data sets and transformation scripts. The system then tracks invocations, annotating the output with information about the runtime environment. The myGrid system [Zha2004], designed to aid biologists in performing computer-based experiments, allows users to model their workflows in a Grid environment. CMCS [Pan2003] is a toolkit for chemists to manage experimental data derived from fields such as combustion research. ESSW [Fre2005] is a data storage system for earth scientists; the system can track data lineage so that errors can be traced, helping maintain the quality of large data sets. Trio [Wid2005] is a data warehouse system that uses data lineage to automatically compute the accuracy of the data. Additional examples can be found in the survey by Bose and Frew [Bos2005].
- **Provenance-aware storage systems.** A provenance-aware storage system supports automatic collection and maintenance of provenance metadata. The system creates provenance metadata as new objects are created in the system and maintains the provenance just as it maintains ordinary file-system metadata. See [PAS]. The Lineage File System [LFS] records the input files, command-line options, and output files when a program is executed; the records are stored in an SQL database that can be queried to reconstruct the lineage of a file.
- **Chain of custody in computer forensics and evidence and change control in software development.** The Vesta [Hey2001] approach uses provenance to make software builds incremental and repeatable.
- **Open Provenance Model.** The Open Provenance Model is a recently proposed abstract data model for capturing provenance. The model aims to make it easier for provenance to be exchanged between systems, to support development of provenance tools, to define a core set of inference rules that support queries on provenance, and to support a technology-neutral digital representation of provenance for any object, regardless of whether or not it is produced by a computer system. See [OPM2007].
- **Pedigree management.** The Pedigree Management and Assessment Framework (PMAF) [SPI2007] enables a publisher of information in a network-centric intelligence gathering and assessment environment to record standard provenance metadata about the source, the manner of collection, and the chain of modification of information as it is passed through processing and assessment.

For further background, see the proceedings of the first USENIX workshop on the theory and practice of provenance [TAP2009].

FUTURE DIRECTIONS

On what categories can we subdivide the topic?

Provenance may be usefully subdivided along three main categories, each of which may be further subdivided, as follows:

- **Representation:** data models and representation structures for provenance (granularity and access control).
- **Management** (creation; access; annotation [mark original documents/resources with provenance metadata]; **editing** [provenance-mark specific fine-grained changes through the life cycle]; **pruning** [delete provenance metadata for performance, security, and privacy reasons]; assurance; and revocation)

- **Presentation** (query [request provenance information]; **present** [display provenance markings]; **alert** [notify when provenance absence, compromise, or fraud is detected])

Other useful dimensions to consider that are cross-cutting with respect to the following dimensions:

- System engineering (human-computer interfaces; workflow implications; and semantic webs)
- Legal, policy, and economic issues (regulation; standards; enforcement; market incentives)

These are summarized in Table 9.1.

What are the major research gaps?

Numerous gaps in provenance and tracking research remain to be filled, requiring a much broader view of the problem space and cross-disciplinary efforts to capture unifying themes and advance the state of the art for the benefit of all communities interested in provenance.

In the following itemization of gaps, the letters R, M, P annotating each point refer to the main categories—**representation**, **management**, and **presentation**, respectively—where uppercase denotes high relevance (R, M, P), and lowercase denotes some relevance (r, m, and p).

- Appropriate definitions and means for manipulating meaningful granularity of information provenance markings. Taxonomy of provenance. (R)
- Given trends in markup languages, the metadata and the underlying data are often intermixed (as in XML), thus presenting challenges in appropriate separation of concerns with data integrity and integrity of the provenance. (R)
- Confidential provenance and anonymous or partially anonymous provenance, to protect sources of information. (R)
- Representing the trustworthiness of provenance. (R)
- Pruning provenance, deleting and sanitizing extraneous item for privacy and purpose of performance. (RMP)
- Efficiently representing provenance. An extreme goal would be to efficiently represent provenance for every bit, enabling bit-grained data transformations, while requiring a minimum of overhead in time and space. (RMp)
- Scale: the need for solutions that scale up and down efficiently. (R)
- Dealing with heterogeneous data types and data sensors, domain specificity, and dependency tracking. (Rm)
- Partial or probabilistic provenance (when the chain of custody cannot be stated with absolute certainty). (RMp)
- Coping with legacy systems. (RM)
- Intrinsic vs. extrinsic provenance and the consistency between them when both are available. (RMp)

TABLE 9.1: Potential Approaches

Category	Definition	Potential Approaches
Representation	Data models and structures for provenance	Varied granularities, integration with access controls
Management	Creation and revocation of indelible distributed provenance	Trustworthy distributed embedding with integrated analysis tools
Presentation	Queries, displays, alerts	Usable human interfaces
System engineering	Secure implementation	Integration into trustworthy systems
Legal, policy, economic issues	Social implications	Regulation, standards, enforcement, incentives

- Developing and adopting tools based on existing research results. (RMP)
- Centralized versus distributed provenance. (M)
- Ensuring the trustworthiness of provenance (integrity through the chain of custody). (M)
- Tracking: where did the information/resources go; how were they used? (M)
- Usable provenance respecting security and privacy concerns. (Mp)
- Information provenance systems should be connected to chain of custody, audit, and data forensic approaches. Provenance should connect and support, not repeat functionality of these related services. (MP)
- User interfaces. When dealing with massive amounts of data from many sources with massive communication processes, how is the end user informed and about what aspects of the information integrity? (P)
- Users of aggregated information need to be able to determine when less reliable information is interspersed with accurate information. It is of critical importance to identify and propagate the source and derivation (or aggregation) of the chain of custody with the information itself. (P)

What are some exemplary problem domains for R&D in this area?

- Computer Emergency Response Teams (CERTs) need to be able to prove from where they got information about vulnerabilities and fixes; when they publish alerts, they should be able to reliably show that the information came from an appropriate, credible source—for example, to avoid publishing an alert based on incorrect information submitted by a competitor. They also need their customers to believe that the information being sent is not from an imposter (although certificates are supposed to take care of this problem).
- Law enforcement forensics for computer-based evidence, surveillance data, and other computer artifacts, of sufficient integrity and oversight to withstand expert counter-testimony.
- Crime statistics and analyses from which patterns of misuse can be deduced.
- Medical and health care information, particularly with respect to data access and data modification.
- Identity-theft and identity-fraud detection and prevention.
- Financial sector—for example, with respect to insider information, funds transfers, and partially anonymous transactions.
- Provenance embedded within digital rights management.

In many of the above examples, some of the provenance may have to be encrypted or anonymized—to protect the identity of sources.

What R&D is evolutionary, and what is more basic, higher risk, game changing?

Information provenance presents a large set of challenges, but significant impact may be made with relatively modest technical progress. For example, it may be possible to develop a coarse-grain information provenance appliance that marks documents traversing an intranet or resting in a data center and makes those markings available to decision makers. Although this imagined appliance may not have visibility into all the inputs used to create a document, it could provide relatively strong assurances about certain aspects of the provenance of the information in question. It is important to find methods to enable incremental rollout of provenance tools and tags in order to maintain compliance with existing practices and standards. Another incremental view is to consider provenance as a static type system for data. Static type systems exist for many programming languages and frameworks that help prevent runtime errors. By analogy, we could create an information provenance system that is able to prevent certain types of misuse of data by comparing the provenance information with policies or requirements.

Resources

With respect to the extensive list of research gaps noted above, resources will be needed for research efforts, experimental testbeds, test and evaluation, and technology transition.

Measures of success

One indicator of success will be the ability to track the provenance of information in large systems that process and transform many different, heterogeneous

types of data. The sheer number of different kinds of sensors and information systems involved and, in particular, the number of legacy systems developed without any attention to maintenance of provenance present major challenges in this domain.

Red Teaming can give added analysis—for example, assessing the difficulty of planting false content and subverting provenance mechanisms.

Also, confidence-level indicators are desirable—for example, assessing the estimated accuracy of the information or the probability that information achieves a certain accuracy level.

More generally, analytic tools can evaluate (measure) metrics for provenance.

Cross-checking provenance with archived file modifications in environments that log changes in detail could also provide measures of success. Efficiency of representations might also be a worthwhile indicator, as would be measures of overhead attributable to maintaining and processing provenance. Metrics that consider human usability of provenance would be very appropriate—especially if they can discern how well people actually are able to distinguish authentic and bogus information based on provenance.

What needs to be in place for test and evaluation?

Testing and evaluating the effectiveness of new provenance systems is challenging because some of the earliest adopters of the technology are likely to be in domains where critical decisions depend on provenance data. Thus, the impact of mistaken provenance could be large.

Potential testbed applications should be considered, such as the following:

- In medical systems, personally identifiable information connected with embarrassing or insurance-relevant information may be used to make life-critical health care decisions.
- An emergency responder system might be considered that could provide more reliable provenance information to decision makers (e.g., who must be evacuated, who has been successfully evacuated from a building).
- A provenance system for the legal profession.
- Credit history and scoring—for example, provenance on credit history data might help reduce delays involved in getting a mortgage despite errors in credit reports.
- Depository services; title history; personnel clearance systems.

References

[Bos2005] R. Bose and J. Frew. Lineage retrieval for scientific data processing: a survey. ACM Computing Surveys, 37(1):1-28, 2005.

[Fos2002] I.T. Foster, J.-S. Voeckler, M. Wilde, and Y. Zhao. Chimera: A virtual data system for representing, querying, and automating data derivation. In Proceedings of the 14th Conference on Scientific and Statistical Database Management, pp. 37-46, 2002.

[Fre2005] J. Frew and R. Bose. Earth System Science Workbench: A data management infrastructure for earth science products. In Proceedings of the 13th Conference on Scientific and Statistical Database Management, p. 180, 2001.

[Hey2001] A. Heydon, R. Levin, T. Mann, and Y. Yu. The Vesta Approach to Software Configuration Management. Technical Report 168, Compaq Systems Research Center, Palo Alto, California, March 2001.

[LFS] Lineage File System (http://theory.stanford.edu/~cao/lineage).

[OPM2007] L. Moreau, J. Freire, J. Futrelle, R.E. McGrath, J. Myers, and P. Paulson. The Open Provenance Model. Technical report, ECS, University of Southampton, 2007 (http://eprints.ecs.soton.ac.uk/14979/).

[Pan03] C. Pancerella et al. Metadata in the collaboratory for multi-scale chemical science. In Proceedings of the 2003 International Conference on Dublin Core and Metadata Applications, 2003.

[PAS] PASS: Provenance-Aware Storage Systems (http://www.eecs.harvard.edu/~syrah/pass/).

[SPI2007] M.M. Gioioso, S.D. McCullough, J.P. Cormier, C. Marceau, and R.A. Joyce. Pedigree management and assessment in a net-centric environment. In Defense Transformation and Net-Centric Systems 2007. Proceedings of the SPIE, 6578:65780H1-H10, 2007.

[TAP2009] First Workshop on the Theory and Practice of Provenance, San Francisco, February 23, 2009 (http://www.usenix.org/events/tapp09/).

[Wid2005] J. Widom. Trio: A system for integrated management of data, accuracy, and lineage. In Proceedings of the Second Biennial Conference on Innovative Data Systems Research, Pacific Grove, California, January 2005.

[Zha2004] J. Zhao, C.A. Goble, R. Stevens, and S. Bechhofer. Semantically linking and browsing provenance logs for e-science. In Proceedings of the 1st International Conference on Semantics of a Networked World, Paris, 2004.

10. Privacy-Aware Security

BACKGROUND

What is the problem being addressed?

The goal of privacy-aware security is to enable users and organizations to better express, protect, and control the confidentiality of their private information, even when they choose to—or are required to—share it with others. **Privacy-aware security** encompasses several distinct but closely related topics, including anonymity, pseudo-anonymity, confidentiality, protection of queries, monitoring, and appropriate accessibility. It is also concerned with protecting the privacy of entities (such as individuals, corporations, government agencies) that need to access private information. This document does not attempt to address the question of what information should be protected or revealed under various circumstances, but it does highlight challenges and approaches to providing technological means for safely controlling access to, and use of, private information. The following are examples of situations that may require limited sharing of private information:

- The need to prove things about oneself (for example, proof of residence)
- Various degrees of anonymity (protection of children online, victims of crime and disease, cash transactions, elections)
- Enabling limited information disclosure sufficient to guarantee security, without divulging more information than necessary
- Identity escrow and management
- Multiparty access controls
- Privacy-protected sharing of security and threat information, as well as audit logs
- Control of secondary reuse
- Remediation of incorrect information that is disclosed, especially if done without any required user approval
- Effective, appropriate access to information for law enforcement and national security
- Medical emergencies (for example, requiring information about allergic reactions to certain medications)

What are the potential threats?

Threats to private information may be intrinsic or extrinsic to computer systems. Intrinsic computer security threats attributable to insiders include mistakes, accidental breach, misconfiguration, and misuse of authorized privileges, as well as insider exploitations of internal security flaws. Intrinsic threats attributable to outsiders (e.g., intruders) include potential exploitations of a wide variety of intrusion techniques. Extrinsic threats arise once information has been viewed by users or made

available to external media (via printers, e-mail, wireless emanations, and so on), and has come primarily outside the purview of authentication, computer access controls, audit trails, and other monitoring on the originating systems.

The central problem in privacy-aware security is the tension between competing goals in the disclosure and use of private information. This document takes no position on what goals should be considered legitimate or how the tension should be resolved. Rather, the goal of research in privacy-aware security is to provide the tools necessary to express and implement trade-offs between competing legitimate goals in the protection and use of private information.

Who are the potential beneficiaries? What are their respective needs?

The beneficiaries for this topic are many and widely varied. They often have directly competing interests. An exhaustive list would be nearly impossible to produce, but some illustrative examples include the following:

- **Individuals** do not generally want to reveal any more private information than absolutely necessary to accomplish a specific goal (transaction, medical treatment, etc.) and want guarantees that the information disclosed will be used only for required and authorized purposes. The ability to detect and correct erroneous data maintained by other organizations (such as credit information bureaus) is also needed.
- **Organizations** do not want proprietary information disclosed for other than specific agreed purposes.
- **Research communities** (e.g., in medical research and social sciences) need access to accurate, specific, and complete data for such purposes as analysis, testing hypotheses, developing potential treatments/solutions.
- **Law enforcement** requires access to personal information to conduct thorough investigations.
- **National security/intelligence** needs to detect and prevent terrorism and hostile activity by nation-states and nonstate actors while maintaining the privacy of U.S. persons and coalition partners.
- **Financial sector** organizations need access to data to analyze for indicators of potential fraud.
- **Health care industries** need access to private patient information for treatment purposes, billing, insurance, and reporting requirements.
- **Product development and marketing** uses data mining to determine trends, identify potential customers, and tune product offerings to customer needs.
- **Business development, partnerships, and collaborations** need to selectively reveal proprietary data to a limited audience for purposes of bidding on a job, engaging in a collaborative venture, pursuing mergers, and the like.
- **Social networks** need means to share personal information within a community while protecting that information from abuse (such as spear-phishing).
- **Governments** need to collect and selectively share information for such purposes as census, disease control, taxation, import/export control, and regulation of commerce.

What is the current state of practice?

Privacy-aware security involves a complex mix of legal, policy, and technological considerations. Work along all these dimensions has struggled to keep up with the pervasive information sharing that cyberspace has enabled. Although the challenges have long been recognized, progress on solutions has been slow, especially on the technology side. At present, there are no widely adopted, uniform frameworks for expressing and enforcing protection requirements for private information while still enabling sharing for legitimate purposes. On the technology side, progress has been made in certain application areas related to privacy. Examples of privacy-enhancing technologies in use today include the following:

- **Access controls** (e.g., discretionary and mandatory, role-based, capability-based, and database management system authorizations) attempt to limit who can access what information,

but they are difficult to configure to achieve desired effects, are often too coarse-grained, and may not map well to actual privacy and data use policies.

- **Encrypted storage and communications** can prevent wholesale loss or exposure of sensitive data but do very little to prevent misuse of data accessed within allowed privileges or within flawed system security.
- **Anonymous credential systems** may enable authorization without necessarily revealing identity (for example, Shibboleth [Shib]).
- **Anonymization techniques**, such as mix networks, onion routing, anonymizing proxy servers, and censorship-resistant access technology, attempt to mask associations between identities and information content.
- **One-time-use technologies**, such as one-time authenticators and smart cards, can also contribute.

At the same time, there are known best practices that, if consistently adopted, would also advance the state of the practice in privacy-preserving information sharing. These include

- Use of trustworthy systems and sound system administration, with strong authentication, differential access controls, and extensive monitoring
- Adherence to the principle of least privilege
- Minimizing data retention time appropriately
- Protecting data in transmission and storage (e.g., with encryption)
- Conducting sensible risk analyses
- Auditing of access audit logs (actually examining them, not just keeping them)
- Privacy policy negotiation and management

What is the status of current research?

Security with privacy appears to require establishment of fundamental trust structures to reflect demands of privacy. It also requires means for reducing the risks of privacy breaches that can occur (accidentally or intentionally) through the use of technologies such as data mining. Ideas for reconciling such technologies in this context include privacy-aware, distributed association-rule mining algorithms that preserve privacy of the individual sites, queries on encrypted data without decrypting, and a new formulation to address the impact of privacy breaches that makes it possible to limit breaches without knowledge of original data distribution.

Digital rights management (DRM) techniques, while not currently applied for privacy protection, could be used to protect information in such diverse settings as health care records and corporate proprietary data, allowing the originator of the information to retain some degree of access control even after the information has been given to third parties, or providing the ability later to identify the misusers. A significant challenge to the DRM approach is the development of an indisputable definition of who controls the distribution. For example, should medical information be controlled by the patient, by doctors, by nurses, by hospitals, or by insurance companies, or by some combination thereof? Each of them may be the originator of different portions of the medical information. Information provenance (Section 9) interacts with privacy in defining the trail of who did what with the medical information, and both interact with system and information integrity.

Many examples of ongoing or planned privacy-related research are of interest here. For example, the following are worth considering. NSF Trustworthy Computing programs have explicitly included privacy in recent solicitations (http://www.nsf.gov/funding/). Some research projects funded by the National Research Council Canada are also relevant (http://iit-iti.nrc-cnrc.gc.ca/r-d/security-securite_e.html), as are British studies of privacy and surveillance, including a technology roadmap (http://www.raeng.org.uk/policy/reports/pdf/dilemmas_of_privacy_and_surveillance_report.pdf).

Other privacy related research includes the following:

- Microsoft Research database privacy: (http://www.research.microsoft.com/jump/50709 and http://www.microsoft.com/mscorp/twc/iappandrsa/research.mspx)

- Project Presidio: collaborative policies and assured information sharing (http://www.projectpresidio.com)
- Stanford University Web Security Research: private information retrieval (http://crypto.stanford.edu/websec/)
- Security with Privacy ISAT briefing (http://www.cs.berkeley.edu/~tygar/papers/ISAT-final-briefing.pdf)
- Naval Research Lab: Reputation in Privacy Enhancing Technologies (http://chacs.nrl.navy.mil/publications/chacs/2002/2002dingledine-cfp02.pdf)
- ITU efforts related to security, privacy, and legislation: (http://www.itu.int/ITU-D/cyb/publications/2006/research-legislation.pdf)
- DHS report on the ADVISE program (http://www.dhs.gov/xlibrary/assets/privacy/privacy_rpt_advise.pdf)
- UMBC Assured privacy preserving data mining, recipient of DoD's MURI award (http://ebiquity.umbc.edu/blogger/tag/muri/)
- Anonymous communication (http://freehaven.net/anonbib)
- Statistics research community, as in the Knowledge Discovery and Data Mining (KDD) conferences (http://sigkdd.org)
- Framework for privacy metrics [Pfi+2001])
- See also (http://www.itaa.org/infosec/faith.pdf) and (http://www.schneier.com/blog/archives/2007/03/security_plus_p.html)

FUTURE DIRECTIONS

On what categories can we subdivide the topic?

For purposes of a research and development roadmap, privacy-aware information sharing can be usefully divided along the following categories, directly mirroring the gaps noted above. See Table 10.1.

- **Selective disclosure and privacy-aware access to data:** theoretical underpinnings and system engineering.
- **Specification frameworks** for providing privacy guarantees: languages for specifying privacy policies, particularly if directly implementable; specifications for violations of privacy; and detecting violations of privacy.
- **Policy issues:** establishing privacy policies, data correction, propagation of updates, privacy implications of data integrity. This also includes **legal** (aspects of current law that constrain technology development; aspects of future law that could enable technology development; questions of jurisdiction), **standards** (best practices; privacy standards analogous to Sarbanes-Oxley; HIPAA), and **economics and security** (e.g., http://www.cl.cam.ac.uk/~rja14/econsec.html).

What are the major research gaps?

Following are some of the gaps in privacy-aware security that need to be addressed.

Selective disclosure and privacy-aware access

- Sound bases are needed for selective disclosure through techniques such as attribute-based encryption, identity-based encryption, collusion-resistant broadcast encryption, private information retrieval (PIR), and oblivious transfer.
- How do we share data sets while reducing the likelihood that arbitrary users can infer individual identification? (The U.S. Census Bureau has long been concerned about this problem.)
- Data sanitization techniques are needed that are nonsubvertible and that at the same time do not render analysis useless.
- More generally, data quality must be maintained for research purposes while protecting privacy, avoiding profiling or temporal analysis to deanonymize source data.
- Irreversible transformations of content are needed that exhibit statistical characteristics

consistent with the original data without revealing the original content.

- Privacy and security for very large data sets does not scale easily—for example, maintaining privacy of individual data elements is difficult.
- Associations of location with users and information may require privacy protection, particularly in mobile devices.
- Low-latency mix networks can provide anonymization, but need further research.
- Mechanisms to enforce retention limits are lacking.
- Sharing of security information such as network trace data needs privacy controls.

Specification frameworks

- Specification frameworks for expressing privacy guarantees are weak or missing. In particular, specification and enforcement of context-dependent policies for data sharing and use are needed.

Policy issues

- Distinctions between individual and group privacy are unclear.
- Release of bogus information about individuals is poorly handled today. However, with stronger protection it becomes more difficult to check validity of information.
- Information gathered from some persons can allow probabilistic inference of information about others.
- Policies for data collection and sharing with regard to privacy are needed, especially relating to what can be done with the private data. For example, who are the stakeholders in genetic information? What policies are needed for retention limits?
- Communications create further privacy problems relating to identification of communication sources, destinations, and patterns that can reveal information, even when other data protections are in place.
- Policies are needed for dealing with privacy violations, detection of violations, consequences of violations, and remediation of damage.

What are some exemplary problems for R&D on this topic?

Several problem domains seem particularly relevant, namely, data mining for medical research, health care records, data mining of search queries, census records, and student records at universities.

What R&D is evolutionary and what is more basic, higher risk, game changing?

Near term

- Deriving requirements for automating privacy policies: learning from P3P
- Policy language development
- Implement best practices
- Research into legal issues in communications privacy

TABLE 10.1: Potential Approaches

Categories	Definition	Potential Approaches
Selective disclosure and privacy-preserving access to data	Technology to support privacy policies	Varied granularities, integration with access controls and encryption
Specification frameworks	Creation and revocation in distributed provenance	Implementable policy languages, analysis tools
Other privacy issues	Policies and procedures to support privacy	Canonical policies, laws, standards, economic models underlying privacy

Medium term

- Anonymous credentials
- Role-based Access Control (RBAC)
- Attribute-based encryption
- Distributed RBAC: no central enforcement mechanism required
- Protection against excess disclosure during inference and accumulation
- Application of DRM techniques for privacy
- Searching encrypted data without revealing the query; more generally, computation on encrypted data

Long term

- Private information retrieval (PIR)
- Multiparty communication
- Use of scale for privacy
- Resistance to active attacks for deanonymizing data
- Developing measures of privacy

Game changing

- Limited data retention
- Any two databases should be capable of being federated without loss of privacy (privacy composability)
- Low-latency private communications resistant to timing attack

Resources

This topic is research-intensive, with considerable needs for testbeds demonstrating effectiveness and for subsequent technology transfer to demonstrate the feasibility of the research. It will require considerable commitment from government funding agencies, corporations, and application communities such as health care to ensure that the research is relevant and that it has adequate testbeds for practical applications. It will also engender considerable scrutiny from the privacy community to ensure that the approaches are adequately privacy preserving.

Measures of success

A goal for addressing concerns regarding both data mining and identity theft is to quantify users' ability to retain control of sensitive information and its dissemination even after it has left their hands. For data mining, quantitative measures of privacy have been proposed only recently and are still fairly primitive. For example, it is difficult to quantify the effect of a release of personal information without knowing the full context with which it may be fused and within which inferences may be drawn. Evaluation and refinement of such metrics are certainly in order.

Useful realistic measures are needed for evaluating privacy and for assessing the relative values of information.

Possible measures of progress/success include the following:

- Rate of publication of privacy-breach stories in the media.
- Database measures: Can we simulate a database without real data? How effective would approaches be that cleanse data by randomization? Can we use such approaches to derive metrics? (Statistical communities have worked on this, as in determining statistical similarity of purposely fuzzed data sets.) How many queries are needed to get to specific data items for individuals in databases that purport to hide such information?
- Adversary work factors to violate privacy.
- Risk analysis: This has been applied to security (albeit somewhat haphazardly). Can risk analysis be effectively applied to privacy?
- Costs for identity-fraud insurance.
- Black market price of stolen identity.

What needs to be in place for test and evaluation?

Access to usable data sets is important, for example,

- Census data (see http://www.fedstats.gov)
- Google Trends
- PREDICT (e.g., network traffic data; http://www.predict.org)
- Medical research data
- E-mail data (e.g., for developing spam filters)

Possible experimental testbeds include the following:

- Isolated networks and their users
- Virtual societies

In addition, privacy Red Teams could be helpful.

References

[Bri+1997] J. Brickell, D.E. Porter, V. Shmatikov, and E. Witchell. Privacy-preserving remote diagnostics, CCS '07, October 29 – November 2, 2007.

[Pfi+2001] A. Pfitzmann and M. Köhntopp. Anonymity, unobservability, and pseudonymity: A proposal for terminology. In *Designing Privacy Enhancing Technologies*, pp. 1-9, Springer, Berlin/Heidelberg, 2001.

[Rab1981] M. Rabin. How to exchange secrets by oblivious transfer. Technical Report TR-81, Aiken Computation Laboratory, Harvard University, 1981.

[Shib] The Shibboleth System (http://shibboleth.internet2.edu/).

Many additional references can be found by browsing the URLs noted above in the text of this section.

11. Usable Security

BACKGROUND

What is the problem being addressed?

Security policy making tends to be reactive in nature, developed in response to an immediate problem rather than planned in advance based on clearly elucidated goals and requirements, as well as thoughtful understanding and analysis of the risks. This reactive approach gives rise to security practices that compromise system usability, which in turn can compromise security—even to the point where intended improvements in a system's security posture are negated. Typically, as the security of systems increases, the usability of those systems tends to decrease, because security enhancements are commonly introduced in ways that are difficult for users to comprehend and that increase the complexity of users' interactions with systems. Any regular and frequent user of the Internet will readily appreciate the challenge of keeping track of dozens of different passwords for dozens of different sites, or keeping up with frequent patches for security vulnerabilities in myriad applications. Many users also are confused by security pop-up dialogs that offer no intuitive explanation of the apparent problem and, moreover, appear completely unable to distinguish normal, legitimate activity, such as reading e-mail from a friend, or from a phishing attempt. Such pop-ups are typically ignored, or else blindly accepted [Sun+09].

People use systems to perform various tasks toward achieving some goal. Unless the tasks at hand are themselves security related, having to think about security interferes with accomplishing the user's main goal. Security as it is typically practiced in today's systems increases complexity of system use, which often causes confusion and frustration for users. When the relationship between security controls and security risks is not clear, users may simply not understand how best to interact with the system to accomplish their main goals while minimizing risk. Even when there is some appreciation of the risks, frustration can lead users to disregard, evade, and disable security controls, thus negating the potential gains of security enhancements.

Security must be usable by persons ranging from nontechnical users to experts and system administrators. Furthermore, systems must be usable while maintaining security. In the absence of usable security, there is ultimately no effective security. The need for usable security and the difficulties inherent in realizing adequate solutions are increasingly being recognized. In attempting to address the challenges of usability and security, several guiding principles are worth considering. Furthermore, when we refer here to usable security, we are really concerned with trustworthy systems whose usability has been designed into them through proactive requirements, constructive architectures, sound system and software development practices, and sensible operation. As observed in previous sections, almost every system component and every step in the development process has the potential to compromise trustworthiness. Poor usability is a huge potential offender.

Security issues must be made as transparent as possible. For example, security mechanisms, policies, and controls must be intuitively clear and perspicuous to all users and appropriate for each user. In particular, the relationships among security controls and security risks must be presented to users in ways that can be understood in the context of system use.

Users must be considered as fundamental components of systems during all phases of the system life cycle. Different assumptions and requirements pertaining to users' interactions with systems must be made explicit to each type of user—novices, intermittent users, experts, and system administrators, to name a few. In general, one-size-fits-all approaches are unlikely to succeed.

Relevant education about security principles and operational constraints must be pervasive. Security issues can never be completely hidden or transparent. There will always be the possibility of conflict between what users might want to accomplish most easily and the security risks involved in doing so. Helping users to understand these trade-offs must be a key component of usable security.

Security metrics must take usability into account. Although one might argue that a system with a certain security control is in principle more secure than an otherwise equivalent system without that control—for example, a web browser that supports client/server authentication vs. one that does not—the real security may in fact be no greater (and possibly even less) in a system that implements that security control, if its introduction compromises usability to the point that users are driven to disable it or switch to an alternative system that is more user friendly but less secure.

What are the potential threats?

The threats from the absence of usable security are pervasive and mostly noted in the above discussion. However, these threats are somewhat different from those in most of the other 10 topics—in that the threats are typically more likely to arise from inactions, inadvertence, and mistakes by legitimate users. On the other hand, threats of misuse by outsiders and insiders similar to those in the other topics can certainly arise as a result of the lack of usability.

Who are the potential beneficiaries? What are their respective needs?

Although the problem of achieving usable security is universal—it affects everyone, and everyone stands to benefit enormously if we successfully address usability as a core aspect of security—it affects different users in different ways, depending on applications, settings, policies, and user roles. The guiding principles may indeed be universal, but as suggested above there is certainly no general one-size-fits-all solution. Examples of different categories of users and ways in which they are affected by problems in usable security are shown in Table 11.1.

What is the current state of practice?

Although the importance of security technology is widely recognized, it is often viewed as a hindrance to productivity. Security is poorly understood by nonexperts, and the consequences of disabled or weakened security controls are often indirect and not immediately felt; and the worst effects may be felt by those not directly involved (e.g., credit card fraud), leading users to question the value of having security technology at all.

At the same time, consciousness of security issues is becoming more widespread, and technology developers are paying increasing attention to security in their products and systems. However, usability in general appears not to be much better understood by software practitioners than security is. This situation makes the problem of usable security even more challenging, since it combines two problems that are difficult to solve individually.

Usability of systems tends to decrease as attempts are made to increase security and, more broadly, trustworthiness. Many current security systems rely on humans performing actions (such as typing passwords) or making decisions (such as whether or not to accept an SSL certificate). For example, one e-mail system requires that users reauthenticate every 8 hours to assure that they are actually the authorized person. This requirement is a direct counter to system usability. For example, some web browsers warn users before any script is run. But users may still browse onto a web server that has scripts on every page, causing pop-up alerts to appear on each page.

Many of the potential impacts of security that is not usable involve increased susceptibility to social-engineering attacks.

TABLE 11.1: Beneficiaries, Challenges, and Needs

Beneficiaries	Challenges	Needs
Nontechnical users	Unfamiliar technology and terminology; security risks unclear	Safe default settings; automated assistance with simple, intuitive explanations when user involvement is required
Occasional users	Changing security landscape; deferred security maintenance (e.g., antivirus updates, software patches) inhibits on-demand system use	Automated, offline system maintenance; automated adaptation of evolving security controls to learned usage patterns
Frequent and expert users	Hidden or inflexible security controls intended for nontechnical users; obtrusive security pop-up dialogs	Security controls that adapt to usage patterns; security control interfaces that remain inconspicuous and unobtrusive, yet readily accessible when needed
Users with special needs (e.g., visual, auditory, motor control challenges)	From a security standpoint, similar to other users, but with added challenges arising from special interface needs	Adaptations of security controls (such as biometrics) that accommodate special needs; for example, fingerprint readers may be unsuitable for users with motor control challenges
System administrators	Configuration and maintenance of systems across different user categories; evolving security threats and policies	Better tools that help automatically configure systems according to organizational policies and user requirements; better tools for monitoring security posture and responding to security incidents
System designers	Lack of security and/or usability emphasis in education and training	Design standards and documented best practices for usable security
System developers	Complexity of adding security and usability requirements into development processes	Integrated development environments (IDEs) that incorporate security and usability dimensions
Policy makers	Difficulty in capturing and expressing security requirements and relating them to organizational workflows	Tools for expressing and evaluating security policies, especially with respect to trade-offs between usability (productivity) and security

This might be an adversary sending an e-mail "this configuration change makes your system more usable" to "this patch must be manually installed". But it also involves attackers who gain the trust of users by helping those users cope with difficult-to-use systems. Thus, resistance to social engineering must be built into systems, and suitable requirements and metrics included from the outset of any system development.

A few illustrative examples from the current state of the practice may help illuminate challenges in usable security and identify some promising directions from which broader lessons may be drawn.

Somewhat positive examples of usable security might include transparent file-system encryption. When first introduced, file encryption technology was cumbersome to configure, even for experts, and imposed significant system overhead. Key management was typically either cumbersome, or reduced to one key or perhaps just a few. Many newer operating systems now offer ready-to-use full-disk encryption out of the box, requiring little more than a password from the user, while imposing no noticeable performance penalty.

Other, more mixed examples illustrate how security technology still falls short in terms of usability:

- **Passwords.** Security pitfalls of poorly implemented password schemes have been extensively documented over the years. When users must resort to writing them on slips of paper or storing them unencrypted on handheld devices, the risk of password exposure may outweigh the increased security of strong passwords. Nevertheless, passwords are often simplistically believed to be a usable security mechanism, and elaborate procedures are promulgated purporting to define sensible password practices (with respect to frequency of changing, not using dictionary words, including nonalphabetic characters, etc.). Tools that help users select good passwords and manage their passwords have been touted to enhance both usability and security. However, to make passwords more effective for stronger security, they must be so long and so complex that users cannot remember them, which seriously compromises usability.
- **Security pop-up dialogs.** No matter how much effort is put into making security controls automated and transparent, there are inevitably situations that require users to make security-related decisions. Today, unfortunately, user involvement appears to be required too often and usually in terms that nontechnical users have difficulty understanding, leading to the frustration effects noted earlier.
- **Mail authentication.** There are mechanisms to authenticate senders of valid e-mails, such as SPF (sender permitted from). DomainKeys Identified Mail (DKIM) is an e-mail authentication technology that allows e-mail recipients to verify whether messages that claim to have been sent from a particular domain actually originated there. It operates transparently for end users and makes it easier to detect possible spam and phishing attacks, both of which often rely on domain spoofing. Some large e-mail service providers now support DKIM.
- **Client-side certificates.** Most web browsers and e-mail applications in widespread use today support user authentication via certificates based on public-key cryptography. However, the technology is not well understood by nonexpert users, and typically the integration of client-side certificate authentication into applications makes the use and management of these certificates opaque and cumbersome for users.
- **The SSL lock icon.** This approach gives the appearance of security, but its limitations are not generally understood. For example, it may be totally spoofed. Its presence or absence may also be ignored.
- "**Web of trust**"-like approaches to certificate trust (e.g., Google, Net Trust). Although this seem to enhance usability, many users may not adequately understand the implications of accepting trust information from systems that may be unknown to those users. They are also unlikely to understand fully what factors might be helpful, harmful, or some of each.
- **CAPTCHA systems.** A CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a challenge-response mechanism intended to ensure that the respondent is a human and not a computer. CAPTCHAs are familiar to most web users as distorted images of words or other character sequences that must be input correctly to gain access to some service (such as a free e-mail account). To make a CAPTCHA effective for distinguishing humans from computers, solving it must be difficult for computers but relatively easy for humans. This balance has proven difficult to achieve, resulting in CAPTCHAs that are either breakable by computers or too difficult for humans. Another challenge is to produce CAPTCHAs that accommodate users with special needs.
- **Not accounting for cultural differences and personal disabilities.** For example, people of one ethnic group tend to have difficulty recognizing different faces of people in other ethnic groups, which

could cause usability differences in authentication. Similarly, CAPTCHAs could be culture dependent. In addition, people with a prosopagnosia disorder have difficulty distinguishing between different people by sight. This would seriously impair their ability to distinguish among different pictorial authenticators and CAPTCHAs.

- **Policies and centralized administration.** Lack of user flexibility is common. On the other hand, it is generally unwise to expect users to make security/usability trade-off evaluations.
- **Federated identity management.** Cross-domain access is complex. Simplistic approaches such as single sign-on can lead to trust violations. Conversely, managing too many passwords is unworkable. More work is needed on access cards such as the CAC system, DoD's Common Access Card, (which combines authentication, encryption of files and e-mail, and key escrow) and other such systems to identify security vulnerabilities. In all such systems, usability is critical.
- **PGP, S/MIME, and other approaches to secure e-mail.** Many past attempts to encapsulate encryption into mail environments have been hindered by the lack of seamless usability.
- **Links.** Phishing, cross-site scripting, and related problems with bogus URLs are laden with risks. URLs may seem to increase usability, but malicious misuse of them can seriously diminish security.
- **Overloading of security attributions** in the context of domain-validation certificates. People tend to trust certificates too much or else are overwhelmed by their presence.
- **Revocation.** Dealing with change is typically difficult, but usability may be impaired when revocation is required. If not carefully designed into systems in advance with usability and understandability in mind, mechanisms for revocation are likely to have unintended consequences.

What is the status of current research?

Following is a brief summary of some current research, along with gaps. For background, see [SOU2008].

- **Usable authentication.** For example, visual passwords and various other authentication approaches exist but need much further work to determine whether they can be used effectively. At present, they are often very difficult to use and seem unlikely to scale well to large numbers of passwords.
- **User security.** Currently funded security-related usability research includes the CMU CyLab Usable Privacy and Security Laboratory (CUPS), and Stanford University work on Web integrity. A list of CUPS projects with descriptions and papers can be found at http://cups.cs.cmu.edu.
- **Ease of administration.** Relatively little research exists in this area. An example of a new direction might be making Tor more usable for administration.
- **Highlighting** important changes to systems (e.g., operating systems, middleware, and applications) that could improve security and usability (rather than just one).
- **Reevaluating** decisions/trade-offs made in past systems. A sense of history in cybersecurity is vital but is too often weak.
- **One Laptop Per Child Bitfrost** security model.
- **Integration of biometrics** with laptops (e.g., fingerprint, facial recognition); this is in practice today, for better or worse. It may be good for administration, but perhaps not so good from the point of view of user understanding.

FUTURE DIRECTIONS

On what categories can we subdivide the topic?

We consider the following three categories as a useful subdivision for formulating a research roadmap for usability and security:

- **Interface design** (I)
- **Science of evaluation for usable security** (E)
- **Tool development** (T)

The following are second-level bins, with descriptors defining their relevance to I, E, and T:

- Principles of usable security; a taxonomy of usable security (E)
- Understanding users and their interactions with security controls (IET)
- Usable authentication and authorization technology (IT)
- Design of usable interfaces for security, with resistance to social engineering (I)
- Development tools that assist in the production of systems that are both more secure and more usable (T)
- Adapting legacy systems
- Building new systems
- Usable security for embedded and mobile devices (IET)
- Evaluation approaches and metrics for usability and security (E)
- User education and familiarization with security issues and technology (IE)
- User feedback, experience (e.g., usability bug reports) (E)
- Security policies (especially, implementation of them) that increase both usability and security (ET)
- Tools for evaluating security policies
- Market creation for usable security technology

What are the major research gaps?

Human-computer interaction (HCI) research has made strides in both designing for and evaluating usability of computer systems. However, only a small fraction of this research has focused on usability specifically as it relates to security. At the same time, security research tends to focus on specific solutions to specific problems, with little or no regard for how those solutions can be made practical and, most importantly, transparent to users and system administrators. To the extent that security practitioners do consider the practical implications of their proposed solutions, the result is often a new or modified user interface component for configuring and controlling the security technology, which does little to address the fundamental problem that most users cannot and do not want to be responsible for understanding and managing security technology; they simply want it to do the right thing and stay out of the way.

In short, usable security is not fundamentally about better user interfaces to manage security technology; rather, it is about evaluating security in the context of tasks and features and of the user, and rearchitecting it to fit into that context.

It is important to note the inherently interdisciplinary nature of usability and security. Security researchers and practitioners cannot simply expect that the HCI experts will fix the usability problem for trustworthy systems. Addressing the problem adequately will require close collaboration between members of the security and usability research communities. One goal is to develop the **science of usability as applied to security**. For example, we need to have ways to evaluate the security of novel approaches and out-of-the-box thinking in usable security.

There is a need to increase knowledge of usability among security practitioners. A common lament in industry is that programmers are too rarely taught how to create secure programs, but even those who do receive such training are unlikely to be taught how to provide both security and usability simultaneously. Just as with security, usability is not a property that can easily be added to existing systems, and it is not a property that one member of a large team can provide for everyone else. The implication is that a large body of designers, programmers, and testers needs to have a much deeper understanding of usability. Adding usability to existing curricula would be a good start but could not be expected to pay dividends for years to come. Methods to increase understanding of usability among software developers already working in industry are equally necessary.

We need to identify a useful framework for discussing usability as it relates to security, such as the following:

- Research on usable security "out of the box" (security transparency).
- Identification of the most useful points in the R&D pipeline at which to involve users in the development of trustworthy systems.
- Research into the question of how to evaluate usability as it relates to security. Here we would expect significant contributions

from HCI research that has already developed methodologies for evaluating usability.

- System architectures that starkly reduce the size and complexity of user interfaces, perhaps by simplifying the interface, hiding the complexity within the interface, providing compatible interfaces for different types of users (such as administrators), or various other strategies, without losing the ability to do what must be done especially in times of system or component failures.
- The ability to reflect physical-world security cues in computer systems.
- Consideration of usability from a data perspective; for example, usability needs can drive collection of data that can lead to security problems (PII as authenticators, for example)

Hard problems

- Usable security on mobile devices
- Usable mutual authentication
- Reusable "clean" abstractions for usable security
- Usable management of access controls
- Usable secure certificate services
- Resistance to social engineering

Other areas we might draw on

- Usability in avionics: reducing the cognitive load on pilots
- Lessons from safety in general, especially warnings science
- Lessons from the automotive industry

What are some exemplary problems for R&D on this topic?

One exemplary problem is protecting users against those who pose as someone else on the Internet. Techniques like certificates have not worked. Alerts from browsers and toolbars and other add-ins about suspicious identities of websites or e-mail addresses do not work, because users either do not understand the alerts or do not bother using the tools. Note that, if used properly, these techniques could be effective. The failure is in their lack of easy usability. The goal here should be not just to find any alternative approach, but rather to find approaches that can work well for ordinary users.

Another exemplary problem is the secure handling of e-mail between an arbitrary sender and an arbitrary receiver in a usable way. Judging from the limited use of encrypted e-mail today, existing approaches are not sufficiently usable. Yet, users are regularly fooled into believing that forged e-mail is actually from the claimed sender. It is only a matter of time before serious problems are encountered because of e-mail traveling across its entire path unencrypted and unauthenticated. For a general discussion on why cryptography is typically not very easily used, see [Whi+1999].

Another possibility is configuring an office environment so that only the people who should have access to sensitive data can actually access it—so that such a configuration can be accomplished by users who understand the effects they want to achieve but are not experts in system administration. In addition, if a user decides to modify the access configuration, how could that be done in a usable way, while achieving only the desired modifications (e.g., not making access to sensitive data either more or less restrictive than intended)?

What R&D is evolutionary and what is more basic, higher risk, game changing?

In the short term, the situation can be significantly improved by R&D that focuses on making security technology work sensibly "out of the box"—ideally with no direct user intervention. More basic, higher-risk, game-changing research would be to identify fundamental system design principles for trustworthy systems that minimize direct user responsibility for trustworthy operation.

Near term

- Informing the security research community on the results obtained in the usable security community on the design and execution of usability studies [Cra+2005]
- Developing a bibliography of best practices and developing a community expectation that security researchers will use them in their work
- Identifying the common characteristics of "good" usable security (and also common characteristics of usability done badly)
- Developing a useful framework

for discussing usability (in the context of security)

- Developing interdisciplinary connections between the security and HCI communities (relates to the first bullet above)
- Identifying ways of involving users in the security technology R&D process

Medium term

- Usable access control mechanisms (such as a usable form of RBAC)
- Usable authentication
- Developing a common framework for evaluating usability and security
- Long term
- Composability of usable components: can we put together good usable components for particular functions and get something usable in the total system?
- Tools, frameworks, and standards for usable security

Resources

Designing and implementing systems with usable security is an enormously challenging problem. It will necessitate embedding requirements for usability in considerable detail throughout the development cycle, reinforced by extensive evaluation of whether it was done adequately. If those requirements are incomplete, it could seriously impair the resulting usability. Thus, significant resources—people, processes, and software development—need to be devoted to this challenge.

Measures of success

Meaningful metrics for usable security must be established, along with generic principles of metrics. These must then be instantiated for specific systems and interfaces. We need to measure whether and to what extent increased usability leads to increased security, and to be able to find "sweet spots" on the usability and security curves. Usable security is not a black-and-white issue. It must also consider returns on investment.

We do not have metrics that allow direct comparison of the usability of two systems (e.g., we cannot say definitively that system A is twice as usable as system B), but we do perhaps have some well-established criteria for what constitutes a good usability evaluation. One possible approach would be to develop a usable solution for one of the exemplar problems and demonstrate both that users understand it and that its adoption reduces the incidence or severity of the associated attack. For example, demonstrate that a better anti-phishing scheme reduces the frequency with which users follow bogus links. Admittedly, this would demonstrate success on only a single problem, but it could be used to show that progress is both possible and demonstrable, something that many people might not otherwise believe is true about usable security.

What needs to be in place for test and evaluation?

Several approaches could help:

- **Test and evaluation** for usability as part of all applicable research in other areas.
- **Guidelines/How-Tos** for usability studies. (See Garfinkel & Cranor [Cra+2005].)
- A "**Usable Security 101**" course, including how to develop and evaluate usable systems.
- **Standardized testbed** for conducting usability studies (perhaps learning from DETER and PlanetLab).
- **Anonymous reporting system** within a repository for usability problems (perhaps learning from the avionics field).

To what extent can we test real systems?

Usability studies need to be based on real systems. They need not be live systems used to conduct actual business, but they need to be real in the sense that they offer the same interfaces and operate in the same environments as such systems.

Usability competitions might be considered (e.g., who can come up with the most usable system for application/function X that satisfies security requirements Y). A possible analogy would be to the challenge of creating a more usable shopping cart. Building test and evaluation into the entire research and development process is essential.

References

[Cra+2005] L.F. Cranor and S. Garfinkel, editors. *Security and Usability: Designing Secure Systems That People Can Use*. O'Reilly Media, Inc., Sebastopol, California, 2005 (http://www.oreilly.com/catalog/securityusability/toc.html).

[Joh2009] Linda Johansson. *Trade-offs between Usability and Security*. Master's thesis in computer science, Linkoping Institute of Technology Department of Electrical Engineering, LiTH-ISY-EX-3165, 2001 (http://www.accenture.com/xdoc/sv/locations/sweden/pdf/Trade-offs%20Between%20Usiability%20and%20Security.pdf).

[SOU2008] Symposium on Usable Privacy and Security. The fourth conference was July 2008 (http://cups.cs.cmu.edu/soups/2008/).

[Sun+09] J. Sunshine, S. Edelman, H. Almuhimedi, N. Atri, and L.F. Cranor. *Crying Wolf: An empirical study of SSL warning effectiveness*. USENIX Security 2009.

[Whi+1999] Alma Whitten and J.D. Tygar. Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In *Proceedings of the 8th USENIX Security Symposium*, Washington, D.C., August 23–26, 1999, pp. 169–184 (http://www.usenix.org/publications/library/proceedings/sec99/whitten.html).

In addition, several other websites might be worth considering.

http://people.ischool.berkeley.edu/~rachna/security_usability.html
http://www.jnd.org/recommended_readings.html
http://gaudior.net/alma/biblio.html
http://www.laptop.org/
http://mpt.net.nz/archive/2008/08/01/free-software-usability

Appendix A

Appendix A. Interdependencies Among Topics

This appendix considers the interdependencies among the 11 topic areas—namely, which topics can benefit from advances in the other topic areas and which topics are most vital to other topics. Although it is in general highly desirable to separate different topic areas in a modular sense with regard to R&D efforts, it is also desirable to explicitly recognize their interdependencies and take advantage of them synergistically wherever possible.

These interdependencies are summarized in Table A.1.

Table A.1: Table of Interdependencies

	Y:													
X: Topic	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**H**	**M**	**L**
1: Scalable Trustworthiness	-	H	H	H	H	H	H	H	H	H	H	10	0	0
2: Enterprise Metrics	M	-	H	H	H	H	H	H	H	H	H	9	1	0
3: Evaluation Life Cycle	H	M	-	H	H	H	H	H	H	M	H	8	2	0
4: Combatting Insider	H	M	M	-	H	M	M	H	M	M	H	4	6	0
5: Combatting Malware	H	M	M	M	-	M	H	H	M	M	H	4	6	0
6: Global ID Management	H	M	M	H	H	-	M	H	H	H	H	7	3	0
7: System Survivability	H	M	M	H	M	M	-	M	M	L	H	3	6	1
8: Situational Awareness	M	M	M	H	H	M	H	-	M	M	H	4	6	0
9: Provenance	M	M	M	M	H	M	M	H	-	H	H	4	6	0
10: Privacy-Aware Security	M	M	L	H	L	H	M	H	M	-	H	4	4	2
11: Usable Security	M	M	M	M	M	M	M	M	M	M	-	0	10	0
H	5	1	2	7	7	4	5	8	4	4	10	*57		
M	5	9	7	3	2	6	5	2	6	5	0		*50	
L	0	0	1	0	1	0	0	0	0	1	0			*3

* Totals for H, M, and L, for both X and Y.

Note: H = high, M = medium, L = low. These are suggestive of the extent to which:
X can contribute to the success of Y.
Y can benefit from progress in X.
Y may in some way depend in the trustworthiness of X.

Almost every topic area has some potential influence and/or dependence on the success of the other topics, as summarized in the table. The extent to which topic X can contribute to topic Y is represented by the letter H, M, or L, which indicate that Topic X can make high, medium, or low contribution to the success of Y. These ratings, of course are very coarse and purely qualitative. On the other hand, any finer-grained ratings are not likely to be useful in this context. The purpose of the table is merely to illustrate the pervasive nature of some relatively strong interdependencies.

A preponderance of H in a *row* indicates that the corresponding row topic is of fundamental importance to other topics. That is, it can contribute strongly to the success of most other topics.

Examples: rows 1 (SCAL: all H), 2 (METR: 9 H), 3 (EVAL: 8 H).

The preponderance of H in a *column* indicates that the corresponding column topic is a primary beneficiary of the other topics.

Examples: columns 11 (USAB: 10 H), 8 (SITU: 8 H), 4 (INSI: 7 H), 5 (MALW: 7 H).

Not surprisingly, the table is not symmetric. However, there are numerous potential synergies here, such as the following:

- **Scalable Trustworthy Systems** (topic 1) is the one topic that can highly enhance all the other topics. However, its success could also derive significant benefits from advances in some of the other topic areas, most obviously including enterprise-level metrics and the system evaluation life cycle (which together could drive the definitions and assessments of trustworthiness), global-scale identity management, system survivability, and usable security, but also including work on combatting insider misuse and combatting malware.
- **Enterprise-Level Metrics (ELMS)** (topic 2) is particularly interesting. It is one topic to which all other topic areas must contribute to some extent, because each other topic area must explicitly include metrics specific to that area. In the other direction of dependence, the mere existence of thorough and well-conceived enterprise-level metrics would drive R&D in the individual topic areas to help them contribute to the satisfaction of the enterprise-level metrics. This can also inspire the composability of the evaluation of topic metrics into the evaluation of the enterprise-level metrics, which is a major research need. The enterprise-level metrics topic area thus interacts bidirectionally with all the other topics, as exhibited by the H entries in that row and the M entries in that column.
- The **System Evaluation Life Cycle** (topic 3) is similar to Enterprise-Level Metrics (topic 2) in this context. It is fundamental to trustworthiness in almost all the other topic areas, but its evolution must also be driven by feedback from those other topics.
- **Combatting Insider Threats** (topic 4) will share some common benefits with **Combatting Malware and Botnets** (topic 5), particularly with respect to the development and systematic use of fine-grained access controls and audit trails. However, note that combatting insider threats can contribute highly (H) to combatting malware, although the reverse contributions may be somewhat less (M). Both of these topics have significant benefits for the other topics. Also, Situational Understanding (topic 8) is fundamental to both, and clearly is relevant to both insider threats and malware. Thus, the potential synergies here will be very important.
- **Global-Scale Identity Management** (topic 6) and **Provenance** (topic 9) can be mutually beneficial: the former can significantly enhance the latter (H), whereas the latter can enhance the former somewhat less (M), although it can increase the assurance of the former.
- **Survivability of Time Critical Systems** (topic 7) is strongly linked with **Scalable Trustworthy Systems** (topic 1), because survivability is one of the fundamental aspects of trustworthiness. In addition, it is particularly relevant to combatting insider threats and malware.

- **Situational Understanding and Attack Attribution** (topic 8) is important throughout.
- **Privacy-Aware Security** (topic 10) is somewhat of an outlier with respect to strong dependence in both directions. It is only moderately dependent on other topics, and most other topics are only moderately dependent on it. Nevertheless, it is a very important and often neglected broad topic area—one that is becoming increasingly important as more applications become heavily dependent on the need for trustworthy computer systems.
- **Usable Security** (topic 11) is fundamental throughout. It can strongly influence the success of almost all the other topics but is also a critical requirement of each of those topics. Generic gains in achieving usability will have enormous impact throughout, in both directions. This is one of many examples of an iterative symbiotic feedback loop, where advances in usability will help other topics, and advances in other topics will help usability.

The low incidence of low-order interdependencies in Table A.1 may at first seem odd. However, it may actually be a testament to the relative importance of each of the 11 topic areas and the mutual synergies among the topics, as well as the inherently holistic nature of trustworthiness [Neu2006], which ultimately requires serious attention to *all* the critical requirements throughout system architecture, system development, and operation. Failure to satisfy any of these requirements can potentially undermine the trustworthiness of entire systems and indeed entire enterprises.

To illustrate the pervasiveness of the interdependencies summarized in Table A.1, we consider the 11 topics, in greater detail. For each topic, we consider first how success in the other topic areas might contribute to that particular topic (that is, represented by the corresponding column of the table), and then consider how success in that particular topic might benefit the other 10 topics (represented by the corresponding rows of the table). These more detailed descriptions are intended to be beneficial for readers who are interested in a particular column or row. They also amplify some of the concepts raised in the 11 sections of this report.

Topic 1: Scalable Trustworthy Systems

We consider first how success in the other topic areas could contribute to scalable trustworthy systems, and then how success in scalable trustworthy systems might benefit the other topic areas.

What capabilities from other topic areas are required or would be particularly desirable for effective progress in this topic area?

Research on the theory and practice of scalable trustworthiness is essential. Although some of that research must result from the pursuit of scalable trustworthy systems per se, research and development experience from the following topics can also contribute to advances in this topic area.

- **Enterprise-level metrics** (that is, measures of trustworthiness that apply to systems and systems of systems as a whole): Evaluation methodologies must allow composability of lower-layer metrics and the resulting evaluations. Formalization of the ways in which metrics and evaluations can compose should contribute to the composability of scalable systems and their ensuing trustworthiness.
- **System evaluation life cycle**: Methodologies for evaluating security should be readily applicable to trustworthy system developments; evaluations must themselves be composable and scalable. Similar to the enterprise-level metrics topic, advances in evaluation methodologies can contribute to the composability of trustworthy systems of systems.
- **Combatting insider threats**: Various advances here could benefit scalable trustworthy systems, including policy development, access control mechanisms and policies, containment and other forms of isolation, compromise-resistant and compromise-resilient operation, and composable metrics and evaluations applicable to insider threats.
- **Combatting malware**: Advances such as those in the previous topic relating to malware detection and prevention can

also contribute, including the existence of contained and confined execution environments (e.g., sandboxing), along with vulnerability analysis tools and composable metrics.

- **Identity management**: Tools for large-scale trust management would enhance scalability and trustworthiness of systems and of systems of systems.
- **System survivability**: Availability models and techniques, self-healing trusted computing bases (TCBs) and subsystems, robustness analysis, composable metrics and evaluations would all be beneficial to scalable trustworthy systems.
- **Situational understanding and attack attribution**: Of considerable interest would be scalable analysis tools. Such tools must scale in several dimensions, including number of system components, types of system components, and attack time scales.
- **Provenance**: The ability of provenance mechanisms and policies to scale cumulatively and iteratively to entire enterprises and federated applications and be maintained under large-scale compositions of components that would enhance scalable trustworthiness overall. Such mechanisms must be tamper resistant, providing abilities for both protection and detection.
- **Privacy-aware security**: Of considerable interest are cryptographic techniques (for example, functional encryption such as attribute-based encryption that is strongly linked with access controls), authentication, and authorization mechanisms that can scale easily into distributed systems, networks, and enterprises, especially if they transcend centralized controls and management.
- **Usable security**: Techniques are needed for building trustworthy systems that are also usable. Thus, any advances in usability can contribute to the development and maintenance of trustworthiness operationally, especially if they can help with scalability.

With respect to prototype systems, systems of systems and enterprises, testbeds and test environments are needed that can be cost-effective and enable timely evaluations, integrated attention to interface design for internal (developer) and external (user) interfaces, and composability with respect to usability metrics. Methods for accurately evaluating large-scale systems in testbeds of limited size would be useful, especially if the methods themselves can scale to larger systems.

How does progress in this area support advances in others?

Overall, this topic area has significant impact on each of the other areas. Scalable composability would contribute directly or indirectly to almost all areas, particularly global identity management, time-critical system survivability, provenance, privacy-aware security, and usability. Usability is an example of two-way interdependence: a system that is not scalable and not trustworthy is likely to be difficult to use; a system that is not readily usable by users and administrators is not likely to be operationally trustworthy. In addition, usability would be mutually reinforcing with evaluation methodologies and global metrics. Other topic areas can benefit with respect to composability and scalability. Metrics must themselves be composable and scalable in order to be extended into enterprise-level metrics. Time-critical systems must compose predictably with other systems. Global-scale identity management, of course, must scale. Usability must compose smoothly.

More detailed technological issues relating to scalable trustworthy systems might address questions such as the following. What fundamental building blocks might be useful for other topic areas, such as insider threats, identity management, and provenance? Can any of these areas, such as usability, use these building blocks composably? Clearly, detailed metrics are needed for trustworthiness, composability, and scalability. Thoroughly documented examples are needed that cut across different topic areas. For example, trustworthy separation kernels, virtual machine monitors, and secure routing represent areas of considerable interest for the future.

Topic 2: Enterprise-Level Metrics (ELMs)

What capabilities from other topic areas are required for effective progress in this topic area?

Each of the other topic areas is expected to define local metrics relevant to its own area. Those local metrics are likely to influence the enterprise-level metrics.

How does progress in this area support advances in others?

Proactive establishment of sensible enterprise-level metrics would naturally tend to drive refinements of the local metrics.

Topic 3: System Evaluation Life Cycle

What capabilities from other topic areas are required for effective progress in this topic area?

Advances in scalability, composability, and overall system trustworthiness are likely to contribute to the development of scalable, composable evaluation methodologies, and suggest some synergistic evolution. Metrics that facilitate evaluation will also contribute significantly.

How does progress in this area support advances in others?

Effective evaluation methodologies can provide major benefits to all the other topics. Otherwise, the absence of such methodologies leaves significant doubts.

Topic 4: Combatting Insider Threats

What capabilities from other topic areas are required for effective progress in this topic area?

Several dependencies on other topic areas are particularly relevant:

- **Scalable trustworthy** systems would help address remote access by logical insiders as well as local access by physical insiders, by virtue of distributed authentication, authorization, and accountability.
- **Situational understanding** and attack attribution must apply to insiders as well as other attackers. This dependency implies that synergy is required between misuse detection systems and the access controls used to minimize insider misuse.
- **Identity management** relates to the accountability aspects of the insider threat, as well as to remote access by insiders.
- **Malware** can be used by insiders or could act as an insider on behalf of an outside actor. Thus, malware prevention can help combat insider threats.
- **Provenance** can also help combat insider threats. For example, strong information provenance can help detect instances where insiders improperly altered critical data.
- **Privacy-aware security** requires knowledge of insiders who were detected in misuse, as well as mechanisms for privacy.

How does progress in this area support advances in others?

- Progress in **combatting insider threats** will support advances in privacy and survivability for time-critical systems, as well as conventional systems. Controls over insider misuse can also help prevent or at least limit the deleterious effects of malware. The prevention aspects are closely related.
- **Life cycle protection** must account for the insider threat.
- **Survivability of systems** can be aided by knowledge of the presence of potential malware or of insiders who may have been detected in potential misuse.

Topic 5: Combatting Malware and Botnets

What capabilities from other topic areas are required for effective progress in this topic area?

Malware is a principal mechanism whereby machines are taken over for botnets. Significant progress in the malware area will go far toward enabling effective botnet mitigation. Economic analysis of adversary markets supports this area, as well as botnet defense, and may provide background intelligence in support of situational understanding.

How does progress in this area support advances in others?

Progress in the area of inherently secure systems that can be thoroughly monitored and audited will benefit other topics, especially situational understanding. Attribution also links this topic to situational understanding. Advances in detection enable malware repositories, which can be mined to identify families and histories of malware, which in turn may make attribution possible.

Collaborative detection may depend on progress in global-scale identity

management, to prevent adversaries from thwarting such an approach through spoofed information.

Progress in security metrics is likely to make it easier to evaluate the effectiveness of proposed solutions to malware problems.

Topic 6: Global-Scale Identity Management

What capabilities from other topic areas are required for effective progress in this topic area?

Scalable trustworthy systems are essential to provide a sound basis for global identity management. **Privacy-aware security** could be highly beneficial. For example, assurance that remote credentials are in fact what they purport to be would help. In addition, analyses, simulations, and data aggregation using real data require strong privacy preservation and some anonymization or sanitization. **Provenance** will be important for increasing the trustworthiness and reputations of remote identities. **Usability** is fundamental, of course for users as well as administrators. **Survivability** of identity management systems will be critical especially, in real-time control and transactional systems.

How does progress in this topic area support advances in others?

Identity management would contribute to the trustworthiness of large-scale networked systems and certainly help in reducing insider misuse, particularly by privileged insiders who are accessing systems remotely. It would also enhance **privacy-preserving security**—for example, because assurances are required whenever there is sharing of identity-laden information. It could simplify security evaluations. It could also reduce the proliferation of malware if identities, credentials, authentication, authorization, and accountability were systematically enforced on objects and other computational entities.

Topic 7: Survivability of Time Critical Systems

What capabilities from other topic areas are required for effective progress in this topic area?

Advances in the development of **scalable trustworthy systems** would have immediate benefits for system **survivability**. Basic advances in **usability** could help enormously in reducing the burdens on system operators and system administrators of survivable systems. Advances in **situational understanding** would also be beneficial in remediating **survivability** failures and compromises.

How does progress in this topic area support advances in others?

Concise and complete requirements for survivability would greatly enhance enterprise-level metrics and contribute to the effectiveness of evaluation methodologies. They would also improve the development of scalable trustworthy systems overall, because of the many commonalities between survivability, security, and reliability.

Topic 8: Situational Understanding and Attack Attribution

What capabilities from other topic areas are required for effective progress in this topic area?

Effective authentication and authorization would make it significantly harder for an attacker to avoid attribution. This depends on progress in **global-scale identity management**.

Subsystems for detecting and **combatting malware** must be designed to enhance situational understanding and attack attribution. Local malware, of course, is a serious problem. However, botnets and the malware that can compromise unsuspecting systems to make them part of botnets are adversarial enablers supporting important classes of attacks for which situational understanding is critical. Attribution in the case of botnets is difficult because the launch points for attacks are themselves victimized machines, and the adversaries are becoming more adept at concealing their control channels and "motherships" (e.g., via encryption, environmental sensing, and fast-flux techniques [ICANN 2008, Holz 2008]).

Advances in **privacy-aware security** (particularly with respect to privacy-aware sharing of security-relevant information) would address many of the hurdles to sharing as considered in this topic area.

The measures of success enumerated below require fundamental advances in metrics definition, collection, and evaluation.

- Synthetic attacks (emulating the best current understanding of adversary tactics) provide some metrics for attribution. Possible metrics include time to detect, how close to the true origin of the attack (adversary and location), and the rate of fast flux

that can be tolerated while still being able to follow the adversary assets.

- We should examine metrics related to human factors to assess effectiveness of presentation approaches.
- We should explore metrics for information sharing—for example, the tradeoff between how much the sharer reveals versus how actionable the community perceives the shared data to be. This issue may touch on sharing marketplaces and reputation systems.
- The current state of metrics with respect to adversary nets and fast flux are not adequately known. We should examine how SANS and similar organizations collect measurement data.

How does progress in this area support advances in others?

For many attack situations of interest, advances in analysis and attack taxonomy would also support malware defense and therefore mitigate botnets. Systems that are intrinsically monitorable and auditable would presumably be easier to defend and less prone to malware.

Advances in attribution to the ultimate attack source would support advances in defense against botnets and other attacks where the immediate launch point of the attack is itself a victimized machine.

This topic and the survivability area are mutually reinforcing. Reaction and mitigation draw on advances in survivability, for example.

Topic 9: Provenance

What capabilities from other topic areas would facilitate progress in this topic area?

Provenance is dependent on most of the other topics and most of the other topics are dependent on provenance, but a few topics have more direct connections. **Global-scale identity management** is required to track authorship as well as chain-of-custody through information processing systems. **Privacy-aware security** is highly relevant to the dissemination of provenance information. Scalable trustworthiness is essential to trustworthy provenance. **Usability** would be important as well.

How does progress in this area support advances in others?

Trustworthy provenance would contribute significantly to **combatting malware** and to **situational understanding**. It could also contribute to **privacy-aware security**. It would provide considerable improvements in system usability overall.

Topic 10: Privacy-Aware Security

What capabilities from other topic areas are required for effective progress in this topic area?

Information provenance is needed for many different privacy mechanisms applied to data. **Scalable trustworthy systems** are needed to ensure the integrity of the privacy mechanisms and policies. **Combatting insider threats** is essential, because otherwise insiders can completely undermine would-be solutions. **Global-scale identity management** is essential for enterprise-wide privacy. **Usability** is essential, because otherwise mechanisms tend to be misused or bypassed and policies tend to be flouted. **Situational understanding and attack attribution**, as well as the ability to combat malware, may be somewhat less important but still can contribute to the detection of privacy violations.

How does progress in this area support advances in others?

Global-scale identity management can benefit—for example, by being shown how to build identity management systems that protect privacy. The **system evaluation life cycle** can benefit from **provenance**. To some extent, this topic can influence requirements for how **scalable trustworthy systems** are designed and developed.

Topic 11: Usable Security

What capabilities from other topic areas are required for effective progress in this topic area?

- **Identity management**: Large-scale identity management systems could solve one of the most vexing security problems users face today—namely, how to establish trust between and among users and systems, particularly within systems and networks that are easy to use by ordinary users and by administrators.
- **Survivability of time-critical systems**: Advances in availability directly enhance usability, especially whenever

manageability of configurations and remediation of potentially dangerous system configurations are included in the design and operation of those systems.

- **Scalable trustworthy systems**: Large-scale systems that are trustworthy must, by the definition of the usability problem, be usable, or they will not be trustworthy, either architecturally or operationally.
- **Provenance**: Automated tools for tracking provenance could enhance usability by reducing the need for users to consider explicitly the source of the information they are dealing with.
- **Privacy-aware security**: As with the other topics, this topic must address usability as a core requirement.
- **Malware**: Technology that neutralizes the threat posed by malware would be of great benefit to usability, since it could eliminate any need for users to think about malware at all.
- **Metrics and evaluation**: The ability to know how well we are doing in making secure systems usable (and usable systems that maintain security) would be useful; a usable system lets you know whether you got things right or wrong.

How does progress in this area support advances in others?

Usability goes hand in hand with the other topic areas; without success in usability, the benefits of progress in the other areas may be diminished. This applies directly to each of the other topic areas, more or less bilaterally. Usability considerations must be addressed pervasively.

Reference

[Neu2006] Peter G. Neumann. Holistic systems. ACM SIGSOFT *Software Engineering Notes* 31(6):4-5, November 2006.

Appendix B

Appendix B. Technology Transfer

This appendix considers approaches for transitioning the results of R&D on the 11 topic areas into deployable systems and into the mainstream of readily available trustworthy systems.

B.1 Introduction

R&D programs, including cyber security R&D, consistently have difficulty in taking the research through a path of development, testing, evaluation, and transition into operational environments. Past experience shows that transition plans developed and applied early in the life cycle of the research program, with probable transition paths for the research products, are effective in achieving successful transfer from research to application and use. It is equally important, however, to acknowledge that these plans are subject to change and must be reviewed often. It is also important to note that different technologies are better suited for different technology transition paths; in some instances, the choice of the transition path will mean success or failure for the ultimate product. Guiding principles for transitioning research products involve lessons learned about the effects of time/schedule, budgets, customer or end-user participation, demonstrations, testing and evaluation, product partnerships, and other factors.

A July 2007 Department of Defense Report to Congress on Technology Transition noted evidence that a chasm exists between the DoD S&T communities focused on demonstration of a component and/or breadboard validation in a relevant environment and acquisition of a system prototype demonstration in an operational environment. DoD is not the only government agency that struggles with technology transition. That chasm, commonly referred to as the *valley of death*, can be bridged only through cooperative efforts and investments by research and development communities as well as acquisition communities.

In order to achieve the full potential of R&D, technology transfer needs to be a key consideration for all R&D investments. This requires the federal government to move past working models in which most R&D programs support only limited operational evaluations/experiments, most R&D program managers consider their job done with final reports, and most research performers consider their job done with publications. Government-funded R&D activities need to focus on the real end goal, namely *technology transfer*, which follows transition. Current R&D Principal Investigators (PIs) and Program Managers (PMs) are not rewarded for technology transfer. Academic PIs are rewarded for publications, not technology transfer. The government R&D community needs to reward government program managers and PIs for transition progress.

There are at least five canonical transition paths for research funded by the Federal Government. These transition paths are affected by the nature of the technology, the intended end-user, participants in the research program, and

other external circumstances. Success in research product transition is often accomplished by the dedication of the program manager through opportunistic channels of demonstration, partnering, and occasional good fortune. However, no single approach is more effective than a proactive technology champion who is allowed the freedom to seek potential utilization of the research product. The five canonical transition paths can be identified simply, as follows:

- Department/Agency direct to Acquisition (Direct)
- Department/Agency to Government Lab (Lab)
- Department/Agency to Industry (Industry)
- Department/Agency to Academia to Industry (Start-up)
- Department/Agency to Open Source Community (Open Source)

Many government agencies and commercial companies use a measure known as a Technology Readiness Level (TRL). The TRL is a term for discussing the maturity of a technology, to assess the maturity of evolving technologies (materials, components, devices, etc.) prior to incorporating that technology into a system or subsystem. Whereas this mechanism is primarily used within the DoD, it can be considered a reasonable guideline for new technologies for almost any department or agency. Table B.1 lists the various technology readiness levels and descriptions from a systems approach for both hardware and software.

B.2 Fundamental Issues for Technology Transition

What are likely effective ways to transfer the technology?

There is no one-size-fits-all approach to technology transfer. Each of the 11 topic areas will have its own special considerations for effective transitioning. For example, effective transitioning will depend to some extent on the relevant customer bases and the specific applications. However, this section considers what might be common to most of the 11 topics. A few issues that are specific to each topic are discussed subsequently.

It will be particularly important that the results (such as new systems, mechanisms, policies, and other approaches) be deployable incrementally, wherever appropriate.

Technologies that are to be deployed on a global scale will require some innovative approaches to licensing and sharing of intellectual property, and serious planning for test, evaluation, and incremental deployment. They will also require extensive commitments to sound system architectures, software engineering disciplines, and commitment to adequate assurance.

Carefully documented worked examples would be enormously helpful, especially if they are scalable. Clearly, the concepts addressed in this document need to become a pervasive part of education and training. To this end, relevant R&D must be explicitly oriented toward real applicability. Furthermore, bringing the concepts discussed in this topic area into the mainstream of education, training, experience, and practice will be essential.

B.3 Topic-Specific Considerations

In this section, certain issues that are specific to each of the 11 topics are considered briefly.

Topic 1: Scalable Trustworthy Systems

Easy scalability, pervasive trustworthiness, and predictable composability all require significant and fundamental changes in how systems are developed, maintained, and operated. Therefore, this topic clearly will require considerable public-private collaboration among government, industry, and academia, with some extraordinary economic, social, and technological forcing functions (see Section B.4). The marketplace has generally failed to adapt to needs for trustworthiness in critical applications.

Topic 2: Enterprise-Level Metrics (ELMs)

This is perhaps a better-mousetrap analogy: if enterprise-level metrics were well developed and able to be readily evaluated (topic 3), we might presume the world would make a beaten path to their door. Such metrics need to be experimentally evaluated and their practical benefits clearly demonstrated, initially in prototype system environments and ultimately in realistic large-scale applications.

Table B1: Typical Technology Readiness Levels

Technology Readiness Level	Description
1. Basic principles observed and reported.	Lowest level of technology readiness. Scientific research begins to be translated into applied research and development. Examples might include paper studies of a technology's basic properties.
2. Technology concept and/or application formulated.	Invention begins. Once basic principles are observed, practical applications can be invented. Applications are speculative and there may be no proof or detailed analysis to support the assumptions. Examples are limited to analytic studies.
3. Analytical and experimental critical function and/or characteristic proof of concept.	Active research and development is initiated. This includes analytical studies and laboratory studies to physically validate analytical predictions of separate elements of the technology. Examples include components that are not yet integrated or representative.
4. Component and/or breadboard validation in laboratory environment.	Basic technological components are integrated to establish that they will work together. This is relatively "low fidelity" compared to the eventual system. Examples include integration of "ad hoc" hardware in the laboratory.
5. Component and/or breadboard validation in relevant environment.	Fidelity of breadboard technology increases significantly. The basic technological components are integrated with reasonably realistic supporting elements so it can be tested in a simulated environment. Examples include "high fidelity" laboratory integration of components.
6. System/subsystem model or prototype demonstration in a relevant environment.	Representative model or prototype system, which is well beyond that of TRL 5, is tested in a relevant environment. Represents a major step up in a technology's demonstrated readiness. Examples include testing a prototype in a high-fidelity laboratory environment or in simulated operational environment.
7. System prototype demonstration in an operational environment.	Prototype near, or at, planned operational system. Represents a major step up from TRL 6, requiring demonstration of an actual system prototype in an operational environment such as an aircraft, vehicle, or space. Examples include testing the prototype in a test bed aircraft.
8. Actual system completed and qualified through test and demonstration.	Technology has been proven to work in its final form and under expected conditions. In almost all cases, this TRL represents the end of true system development. Examples include developmental test and evaluation of the system in its intended weapon system to determine if it meets design specifications.
9. Actual system proven through successful mission operations.	Actual application of the technology in its final form and under mission conditions, such as those encountered in operational test and evaluation. Examples include using the system under operational mission conditions.

Topic 3: System Evaluation Life Cycle

Similarly, if effective evaluation methodologies could be developed, their usefulness would need to be clearly demonstrated on real systems, as in topic 2. Thoroughly specified and relatively complete requirements would also be required. Given a few well-documented demonstrations of effectiveness, the incentives for technology transfer would be greatly increased.

Topic 4: Combatting Insider Threats

Once again, the proof is in the pudding. Demonstrations of the effectiveness of approaches that combat insider misuse would encourage adoption of the techniques.

Topic 5: Combatting Malware and Botnets

As noted in Appendix A, the commonalities among insider threats and malware suggest that demonstrations of the effectiveness of approaches that combat malware are likely to be rapidly and widely adopted in practice.

Topic 6: Global-Scale Identity Management

It will be important to design mechanisms and policies that can be incrementally deployed. Technologies that are to be deployed on a global scale will require some innovative approaches to licensing and sharing intellectual properties, and serious planning for test, evaluation, and incremental deployment.

Topic 7: Survivability of Time Critical Systems

R&D communities have long understood how to take advantage of fault-tolerance mechanisms. However, system survivability requires an overarching commitment to system trustworthiness that must transcend what has been done in the past.

Topic 8: Situational Understanding and Attack Attribution

R&D in this area has been slow to find its way into commercial products. Recognition of the pervasive needs for monitoring and accountability would be of great value.

Topic 9: Provenance

Provenance would be very useful in finance, government, health care, and many other application areas, and would facilitate forensics.

Topic 10: Privacy-Aware Security

Advances in this topic could be particularly useful in many application areas, such as health care, financial records, communication logs, and so on.

Topic 11: Usable Security

Almost anything that significantly increased the usability of security and helped manage its inherent complexity would be likely to find its way into practice fairly readily.

B.4 Forcing Functions (Some Illustrative Examples)

For several of the 11 topics, this section addresses the question **What are the appropriate roles for government, academia, industry, and markets?** Many of the suggested forcing functions are applicable in other topics as well.

Topic 1: Scalable Trustworthy Systems

The federal government needs to encourage and fund research and development relating to all of the topics considered here, with particular emphasis on trustworthy systems, composability, scalability, and evolutionary system architectures. It also needs to encourage the incorporation of source-available and nonproprietary systems that can demonstrably contribute to trustworthiness.

Academic research needs to pursue theories and supporting tools that enable systematic development of composable and scalable trustworthy systems and all the other topics discussed here.

Commercial developers need to instill a more proactive discipline of principled system developments that allow interoperability among different systems and subsystems, that employ much better software engineering practices, that result in trustworthy systems that are more composable and scalable, and that provide cost-effective approaches for all the topics discussed here.

Topic 4: Combatting Insider Threats

Governments need to establish baselines and standards. Legal issues relating to trap-based defensive strategies and entrapment law should be addressed. Applying these to the many real situations in government activity where insider behavior is a genuine threat would be beneficial. Current government efforts to standardize on authentication and authorization (e.g., the Common Access Card) are worthwhile despite their potential limitations, particularly in helping combat insider misuse. Academia needs to pursue R&D that is realistically relevant to the insider threat. Industry research needs

to be more closely allied with the needs of practical systems with fine-grained access controls and monitoring facilities. Industry is also the most likely source of data sets that contain instances of insider misbehavior, or at least more detailed knowledge of some kind on how real insider misbehavior tends to manifest itself. The marketplace needs to be responsive to customers demanding better system solutions. Note also the possible relevance of HSPD-12 PIV-I and PIV-II.

Various incentive structures might be considered:

- Business cases as incentive (investment vs. potential cost)
- Insurance as financial protection against insiders
- Major players in the bonding markets, who might possibly provide data for research in exchange for better loss-reduction approaches
- Nonfinancial incentives, as in the FAA near-miss reporting, granting some sort of immunity (but being careful not to shoot the whistle-blowers)
- International efforts might include bilateral and multilateral quid-pro-quo cooperations.

Topic 6: Global-Scale Identity Management

Governments need to unify some of the conflicting requirements relating to identity management, credentials, and privacy. The U.S. government needs to eat its own dog food, establishing sound identity management mechanisms and policies, and adhering to them.

Academia needs to recognize more widely the realistic problems of global identity management and to embed more holistic and realistic approaches into research.

Industry needs to recognize the enormous need for interoperability within multivendor and multinational federated systems.

The marketplace needs to anticipate long-term needs and somehow inspire governments, academia, and industry to realize the importance of realistic approaches.

Topic 11: Usable Security

Government

Remove impediments to usability research. For example, federal law currently requires review before data can be used in an experiment/study; simply having the data in your possession does not give you the right to use it (e.g., e-mail you have received and wish to use to test a new spam filtering algorithm); Minimize administrative burdens; making sure Institutional Review Boards (IRBs) are familiar with the unique aspects of usable security research (especially as contrasted, for example medical research); and create mechanisms for expediting usability research approval.

- Avoid inappropriate restrictions that prevent government entities from participating in research.
- Provide suitable funding for basic research in usable security.
- Encourage interdisciplinary research in usable security.
- Adopt usability reviews for security research.
- Establish appropriate standards, criteria, and best practices.
- Pervasively embed usability requirements into the procurement process.
- Reconsider security policies from a usability perspective.
- Ensure that usable security is a criteria for evaluating NSA centers of academic excellence. (This will provide an incentive to get usability into the curriculum.)

Academia

- Incorporate usability pervasively into computer system curricula.
- Lead by example by making their own systems more usably secure.
- Incorporate usability into the research culture by demanding that system security research papers and proposals always address issues of usability.

Industry

- Develop standards for usable security.
- Develop consistent terminology.
- Identify best practices.
- Contribute deployment experience. (Provide feedback to the research community: what works and what does not.)

Appendix C

Appendix C. List of Participants in the Roadmap Development

We are very grateful to many people who contributed to the development of this roadmap for cybersecurity research, development, test, and evaluation. Everyone who participated in at least one of the five workshops is listed here.

Deb Agarwal
Tom Anderson
Paul Barford
Steven M. Bellovin
Terry Benzel
Gary Bridges
KC Claffy
Ben Cook
Lorrie Cranor
Rob Cunningham
David Dagon
Claudiu Danilov
Steve Dawson
Drew Dean
Jeremy Epstein
Sonia Fahmy
Rich Feiertag
Stefano Foresti
Deb Frincke
Simson Garfinkel
Mark Graff
Josh Grosh
Minaxi Gupta
Tom Haigh
Carl Hauser
Jeri Hessman
James Horning
James Hughes
Bob Hutchinson
Cynthia Irvine
Markus Jakobsson
David Jevans
Richard Kemmerer
Carl Landwehr
Karl Levitt
Jun Li
Pat Lincoln
Ulf Lindqvist
Teresa Lunt
Doug Maughan
Jenny McNeill
Miles McQueen
Wayne Meitzler
Jennifer Mekis
Jelena Mirkovic
Ilya Mironov
John Mitchell
John Muir
Deirdre Mulligan
Clifford Neuman
Peter Neumann
David Nicol
Chris Papadopoulos
Vern Paxson
Peter Reiher
Robin Roy
William H. Sanders
Mark Schertler
Fred Schneider
Kent Seamons
John Sebes
Frederick T. Sheldon
Ben Shneiderman
Pete Sholander
Robert Simson
Dawn Song
Joe St Sauver
Sal Stolfo
Paul Syverson
Kevin Thompson
Gene Tsudik
Zach Tudor
Al Valdes
Jamie Van Randwyk
Jim Waldo
Nick Weaver
Rick Wesson
Greg Wigton
Bill Woodcock
Bill Worley
Stephen Yau
Mary Ellen Zurko

Appendix D

Appendix D. Acronyms

A/V	antivirus
AMI	Advanced Metering Infrastructure
BGP	Border Gateway Protocol
C2	command and control
CAC	Common Access Card
CAPTCHA	Completely Automated Public Turing test to tell Computers and Humans Apart
CASSEE	computer automated secure software engineering environment
CERTs	Computer Emergency Response Teams
CMCS	Collaboratory for Multi-scale Chemical Science
COTS	commercial off-the-shelf
CUI	Controlled Unclassified Information
CVS	Concurrent Versions System
DAC	discretionary access controls
DARPA	Defense Advanced Research Projects Agency
DDoS	distributed denial of service
DETER	cyber-DEfense Technology Experimental Research
DHS	Department of Homeland Security
DKIM	DomainKeys Identified Mail
DNS	Domain Name System
DNSSEC	DNS Security Extensions
DoS	denial of service
DRM	digital rights management
ESSW	Earth System Science Workbench
EU	European Union
FIPS	Federal Information Processing Standards
FISMA	Federal Information Security Management Act
GPS	Global Positioning System
HDM	Hierarchical Development Methodology
HIPAA	Health Insurance Portability and Accountability Act
HSI	human-system interaction
HVM	hardware virtual machine
I&A	identification and authentication
I3P	Institute for Information Infrastructure Protection
IDA	Institute for Defense Analyses
IDE	integrated development environment
IDS	intrusion detection system
INL	Idaho National Laboratory
IPS	intrusion prevention system

IPsec	Internet Protocol Security
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
IRB	institutional review board
ISP	Internet service provider
IT	information technology
LPWA	Lucent Personalized Web Assistant
MAC	mandatory access controls
MIT	Massachusetts Institute of Technology
MLS	multilevel security
MTBF	mean time between failures
NIST	National Institute of Standards and Technology
NOC	network operations center
OODA	Observe, Orient, Decide, Act
OS	operating system
OTP	one-time password
P2P	peer-to-peer
P3P	Platform for Privacy Preferences
PDA	personal digital assistant
PGP	Pretty Good Privacy
PII	personally identifiable information
PIR	private information retrieval
PKI	public key infrastructure
PL	programming language
PMAF	Pedigree Management and Assessment Framework
PSOS	Provably Secure Operating System
PREDICT	Protected Repository for the Defense of Infrastructure against Cyber Threats
QoP	Quality of Protection
RBAC	role-based access control
RBN	Russian Business Network
RFID	radio frequency identification
ROM	read-only memory
SBU	Sensitive But Unclassified
SCADA	Supervisory Control and Data Acquisition
SCAP	Security Content Automation Protocol
SIEM	security information and event management
SOHO	small office/home office
SPF	sender permitted from
SQL	Structured Query Language

SRS	Self-Regenerative Systems
SSL	Secure Sockets Layer
T&E	test and evaluation
TCB	trusted computing base
TCP/IP	Transmission Control Protocol/Internet Protocol
TLD	top-level domain
TPM	Trusted Platform Module
TSoS	trustworthy systems of systems
UI	user interface
UIUC	University of Chicago at Urbana-Champaign
USB	universal serial bus
US-CERT	United States Computer Emergency Readiness Team
VM	virtual machine
VMM	Virtual Machine Monitor